

# Properties of $z \sim 3$ to $z \sim 6$ Lyman Break Galaxies. I. Impact of nebular emission at high redshift

S. de Barros<sup>1</sup>, D. Schaerer<sup>1,2</sup>, and D.P. Stark<sup>3</sup>

<sup>1</sup> Geneva Observatory, University of Geneva, 51, Ch. des Maillettes, CH-1290 Versoix, Switzerland  
e-mail: stephane.debarros@unige.ch

<sup>2</sup> CNRS, IRAP, 14 Avenue E. Belin, 31400 Toulouse, France

<sup>3</sup> Kavli Institute of Cosmology and Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB30HA, UK

Received; accepted

## ABSTRACT

**Context.** To gain insight on the mass assembly and place constraints on the star formation history (SFH) of Lyman break galaxies (LBGs) it is important to determine properly and accurately their properties.

**Aims.** We estimate how nebular emission and different SFHs affect parameter estimation and uncertainties.

**Methods.** We present a homogeneous, detailed analysis of the spectral energy distribution (SED) of  $\sim 1700$  LBGs from the GOODS-MUSIC catalog with deep multi-wavelength photometry from  $U$  band to  $8 \mu\text{m}$ , to determine stellar mass, age, dust attenuation and star formation rate. Using our SED fitting tool which takes into account nebular emission, we explore a wide parameter space. We also explore a set of different star formation histories.

**Results.** Nebular emission is found to significantly affect the determination of the physical parameters for the majority of LBGs at redshift  $z \sim 3-6$ . We identify two populations of galaxies by determining the importance of the contribution of emission lines. We find that  $\sim 65\%$  of LBGs show detectable signs of emission lines, whereas  $\sim 35\%$  show weak or no emission lines. This distribution is found over the entire redshift range. We interpret these groups as actively star forming and more quiescent LBGs respectively. We find that models with constant star formation cannot reproduce the entire range of observed colors, whereas models with nebular emission and variable (declining or rising) star formation histories succeed. Other arguments favoring episodic star formation and relatively short star formation timescales are also discussed. Taking into account nebular emission generally leads, for a given SFH, to a younger age, lower stellar mass, higher dust attenuation, and higher star formation rate, although with increased uncertainties. We find a trend of increasing dust attenuation with galaxy mass, and a large scatter in the  $\text{SFR}-M_\star$  relation. Our analysis yields a trend of increasing specific star formation rate with redshift, as predicted by recent galaxy evolution models.

**Conclusions.** The physical parameters of approximately two thirds of high redshift galaxies are significantly modified when nebular emission is taken into account. SED models including nebular emission shed new light on the properties of LBGs with numerous important implications.

**Key words.** galaxies: starburst – galaxies: high redshift – galaxies: evolution – galaxies: star formation

## 1. Introduction

The study of star formation history at high redshift is an important issue, with both intensive observational (e.g. Bouwens et al. 2009; Stark et al. 2009; Bouwens et al. 2011) and theoretical work (e.g. Finlator et al. 2007; Bouché et al. 2010; Finlator et al. 2011; Weinmann et al. 2011). While theoretical studies attempt to reproduce observed trends, as UV luminosity function, the lack of available spectroscopic data at  $z > 3$  leads to rely on parameter estimation based on indirect indicators as UV  $\beta$  slope for dust attenuation or UV luminosity for star formation rate (e.g. Bouwens et al. 2011) or using SED fitting (e.g. Shapley et al. 2001; Papovich et al. 2001; Verma et al. 2007; Stark et al. 2009; Lee et al. 2011). The ability of SED fitting to estimate physical parameters has been considerably improved with large multi-wavelength surveys, thanks to the commissioning of new space and ground-based facilities in the last decades as the *Hubble Space Telescope*, *Spitzer* or the *Very Large Telescope*. Indeed, an accurate determination of physical parameters is necessary to put strong constraints on stellar mass assembly and evolution of physical parameters.

Although nebular emission (i.e. emission lines and nebular continuous emission from HII regions) is ubiquitous in regions

of massive star formation, strong or dominant in optical spectra of nearby star forming galaxies, and present in numerous types of galaxies, its impact on the determination of physical parameters of galaxies, in particular at high redshift, has been neglected until recently (cf. overview in Schaerer & de Barros 2011). Several spectral models of galaxies have indeed included nebular emission in the past (e.g. Charlot & Longhetti 2001; Fioc & Rocca-Volmerange 1997; Anders & Fritze-v. Alvensleben 2003; Zackrisson et al. 2008); however they have not been applied to the analysis of distant galaxies.

Zackrisson et al. (2008) showed that nebular emission can significantly affect broad-band photometry. Schaerer & de Barros (2009) included for the first time nebular emission to fit SED of a sample of Lyman break galaxies at  $z \sim 6$ , and showed that nebular lines strongly affect age estimation, since some lines can mimic a Balmer break. Ages are strongly reduced, which can lead to reconsider star formation rate estimations from UV luminosity, since the standard relation used to convert UV luminosity into SFR is based on the assumption of a constant star formation activity during 100 Myr (Kennicutt 1998; Madau et al. 1998). The analysis of a sample of  $z \sim 6-8$  LBGs observed with HST and Spitzer further demonstrated the potential impact of

nebular emission on the physical parameters derived from SED fits of high- $z$  galaxies (Schaerer & de Barros 2010).

It has now become clear (Schaerer & de Barros 2009, 2010; Ono et al. 2010; Lidman et al. 2012) that nebular emission (both lines and continuum emission) must be taken into account for the interpretation of photometric measurements of the SEDs of star-forming galaxies such as Lyman-alpha emitters and Lyman break galaxies – the dominant galaxy populations at high- $z$ . Furthermore, as testified by the presence of  $\text{Ly}\alpha$  emission, a large and growing fraction of the currently known population of star-forming galaxies at high redshift shows emission lines (Ouchi et al. 2008; Stark et al. 2010; Schaefer et al. 2011; Curtis-Lake et al. 2012; Schenker et al. 2012).

In parallel a lot of diverse evidence for galaxies with strong emission lines and/or strong contributions of nebular emission to broad-band filters has been found at different redshifts, e.g. by Shim et al. (2011); McLinden et al. (2011); Atek et al. (2011); Trump et al. (2011); van der Wel et al. (2011). It is therefore obvious that a systematic study of the impact of nebular emission on the physical parameters of LBGs at different redshift is urgently needed. This is the main objective of the present work.

Several other studies included the effect of nebular emission (e.g. Ono et al. 2010; Schaefer & de Barros 2010; Acquaviva et al. 2011; McLure et al. 2011), providing further evidences of the impact of nebular emission on physical parameter estimation, leading to lower stellar masses and higher SFRs, possibly reconciling expected evolution of increasing specific star formation rate (sSFR) with redshift from hydrodynamical simulations, with sSFR determined from observations (Bouché et al. 2010). Indeed, studies which do not consider nebular emission found no evolution of sSFR with redshift above  $z \sim 2$  (e.g. Stark et al. 2009; González et al. 2010).

Another issue is that UV  $\beta$  slope seems to indicate that galaxies at high redshift are essentially dust free (Bouwens et al. 2011), while other studies provide evidences for some high redshift obscured galaxies (Verma et al. 2007; Yabe et al. 2009; Schaefer & de Barros 2010). Furthermore, there is a possible trend of increasing dust attenuation with the stellar mass (Schaerer & de Barros 2010), trend already established at lower redshift (Brinchmann et al. 2004).

Stark et al. (2009) was the first attempt of constraining star formation history by studying evolution of LBGs sample uniformly selected among different bins of redshift. In this work, we use a similar approach, with a large sample of LBGs covering four bins of redshift between  $z \sim 3$  to  $z \sim 6$ , using an up-to-date photometric redshift and SED-fitting tool that treats the effects of nebular emission on the SEDs of galaxies. This homogeneous analysis provides the main physical parameters, as star formation rate, stellar mass, age and reddening.

Our paper is structured as follows. The observational data are described in Sect. 2, and the method used for SED modelling is described in Sect. 3. The results are presented in Sect. 4 and discussed in Sect. 5. Section 6 summarises our main conclusions. We adopt a  $\Lambda$ -CDM cosmological model with  $H_0=70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m=0.3$  and  $\Omega_\Lambda=0.7$ . All magnitudes are expressed in the AB system (Oke & Gunn 1983)

## 2. Data

### 2.1. The GOODS Fields

We focus our analysis on the data from the Great Observatories Origins Deep Survey (GOODS). Detailed descriptions of the datasets are available in the literature (Giavalisco et al. 2004;

Santini et al. 2009), so we only provide a brief summary here. The GOODS-S and GOODS-N survey areas each cover roughly  $160 \text{ arcmin}^2$  and are centered on the *Chandra* Deep Field South (CDF-S; Giacomini et al. 2002) and the *Hubble* Deep Field North (HDF-N; Williams et al. 1996). Extensive multiwavelength observations have been conducted in each of these fields. In this paper, we utilize optical imaging from the Advanced Camera for Surveys (ACS) onboard the Hubble Space Telescope (HST). Observations with ACS were conducted in F435W, F606W, F775W, and F850LP (hereafter  $B_{435}$ ,  $V_{606}$ ,  $i'_{775}$ ,  $z'_{850}$ ) toward GOODS-S and GOODS-N (Giavalisco et al. 2004). The average  $5\sigma$  limiting magnitudes in the v2 GOODS ACS data ( $0''.35$  diameter photometric aperture) are  $B_{435}=29.04$ ,  $V_{606}=29.52$ ,  $i'_{775}=29.19$ , and  $z'_{850}=28.54$ . We also make use of U and R-band observations of GOODS-S taken with the ESO Very Large Telescope (VLT) with the VIMOS wide field imager (Le Fèvre et al. 2003), provided by Nonino et al. (2009), with a  $1\sigma$  limiting magnitude ( $1''$  radius aperture) reaching  $U \approx 29.8$ .

In the near-infrared, we utilize publicly available deep J, H and K-band observations of GOODS-S (PI: C. Cesarsky) using the ISAAC camera on the VLT. The sensitivities vary across the field depending on the effective integration time and seeing FWHM. Average  $5\sigma$  magnitude limits (corrected for the amount of flux that falls outside of the  $1''.0$  diameter aperture) are  $J \approx 25.2$ ,  $H \approx 24.7$  and  $K_s \approx 24.7$ .

Deep *Spitzer* imaging is available toward both GOODS fields with the Infrared Array Camera (IRAC) as part of the “Super Deep” Legacy program (Dickinson et al. *in prep*). Details of the observations have been described in detail elsewhere (Eyles et al. 2005; Yan et al. 2005; Stark et al. 2007) so we do not discuss them further here. The  $5\sigma$  limiting magnitudes of the IRAC imaging are  $\approx 26.3$  at  $3.6\mu\text{m}$  and  $\approx 25.9$  at  $4.5\mu\text{m}$  using  $2''.4$  diameter apertures and applying an aperture correction.

Optical, near and mid-infrared photometry are described in Santini et al. (2009). Non-detections are included in the SED fit with *Hyperz* by setting the flux in the corresponding filter to zero, and the error to the  $1\sigma$  upper limit.

### 2.2. Dropout selection

Galaxies at  $z \approx 3, 4, 5$ , and  $6$  are selected via the presence of the Lyman-break as it is redshifted through the U,  $B_{435}$ ,  $V_{606}$ , and  $i'_{775}$  bandpasses, respectively. Selection of Lyman break galaxies at these redshifts has now become routine (Stanway et al. 2003; Giavalisco et al. 2004; Bunker et al. 2004; Beckwith et al. 2006; Bouwens et al. 2007). In order to ensure a consistent comparison of our samples to these previous samples, we adopt color criteria which are similar to those used in Beckwith et al. (2006), and very similar to those used by Bouwens et al. (2007). These criteria have been developed to select galaxies in the chosen redshift interval while minimizing contamination from red galaxies likely to be at low redshift.

As mentioned in the previous section, galaxies are selected from the GOODS-MUSIC catalog v2 (Santini et al. 2009). It contains 14 999 objects selected in either the  $z_{850}$  band or the  $K_s$  band or at  $4.5\mu\text{m}$ .

The selection dropout criteria used in this work are identical to those from Nonino et al. (2009) and Stark et al. (2009), with an additional constraint for U dropout ( $S/N(U) < 2$ ). This latter criteria helps greatly in removing contaminants, but it also removes galaxies at the low redshift tail of the U-drop selection. This selection leave us 440 U drops, 859 B drops, 277 V drops and 66 i drops.

### 3. Method

#### 3.1. SED fitting tool

We use a recent, modified version of the Hyperz photometric redshift code of Bolzonella et al. (2000), taking into account nebular emission (lines and continua). We consider a large set of spectral templates (Bruzual & Charlot 2003), covering different metallicities and a wide range of star formation (SF) histories (exponentially decreasing, constant and rising SF), and we add the effects of nebular emission following our method presented in Schaerer & de Barros (2009, 2010). We account for attenuation from the intergalactic and the interstellar medium and varying redshift. With these assumptions we fit the observed SEDs by straightforward least-square minimization.

In practice we adopt a spectral templates computed for a Salpeter IMF (Salpeter 1955) from 0.1 to 100  $M_{\odot}$ , and we properly treat the returned ISM mass from stars. Nebular emission from continuum processes and lines is added to the spectra predicted from the GALAXEV models as described in Schaerer & de Barros (2009), proportionally to the Lyman continuum photon production. The relative line intensities of He and metals are taken from Anders & Fritze-v. Alvensleben (2003), including galaxies grouped in three metallicity intervals covering  $\sim 1/50$ – $1 Z_{\odot}$ . Hydrogen lines from the Lyman to the Brackett series are included with relative intensities given by case B. For galactic attenuation we use the Calzetti law (Calzetti et al. 2000). The IGM is treated following Madau (1995).

To examine the effects of different star formation histories (SFH) and for comparison with other studies we define three sets of models:

- *Reference model* (REF): constant star formation rate, minimum age  $t > 50$  Myr, and solar metallicity.
- *Decreasing model* (DEC): exponentially declining star formation histories ( $\text{SFR} \propto \exp(-t/\tau)$ ) with variable timescales  $\tau$ . Metallicity and  $\tau$  are free parameters.
- *Rising model* (RIS): rising star formation rate. We use the mean rising star-formation history from the simulations of Finlator et al. (2011, their Fig. 1). To describe this case, we assume that SF starts at age=0 and grows by 2.5 dex during 0.8 Gyr following their functional dependence. After this period we set  $\text{SFR} = 0$ . Metallicity is a free parameter.

Furthermore, we define three options concerning the treatment of nebular emission:

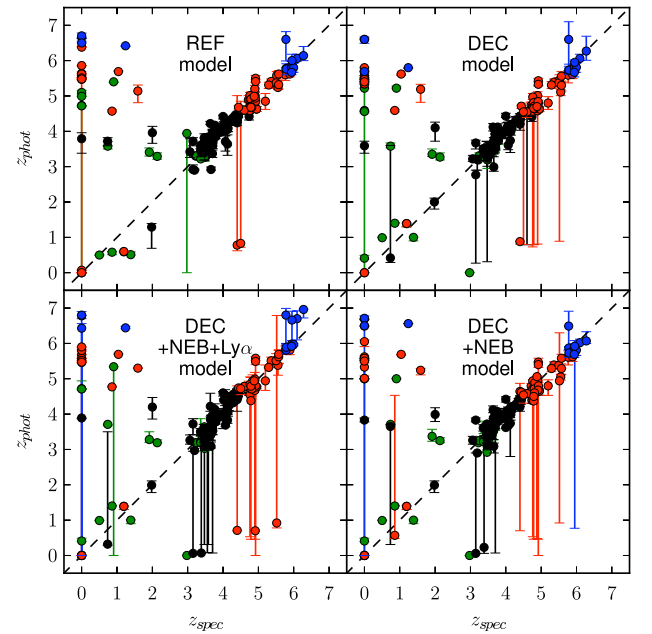
- No nebular emission.
- +NEB: including nebular continuum emission and lines except  $\text{Ly}\alpha$ , since this line may be attenuated by radiation transfer processes inside the galaxy or by the intervening intergalactic medium.
- +NEB+ $\text{Ly}\alpha$ : including nebular emission (all lines and continuum processes).

SEDs fits to all galaxies have been computed for each of the above model sets and nebular options, i.e. for nine different combinations. This allows us to examine in detail the impact these assumptions/options have on the derived physical parameters, and to compare also their fit quality. For the B-drop sample we have also tested the effect of the SMC extinction law of Prevot et al. (1984).

In general, the free parameters of the SED fits are: the redshift  $z$ , metallicity  $Z$  (of stars and gas), the star formation history parametrised by  $\tau$ , the age  $t$  defined since the onset of star-formation, and the attenuation  $A_V$ . For the reference model set

the SFH and metallicity are fixed and the age limited to a minimum. For the RIS model set, the SFH is also fixed. In all cases, we consider  $z \in [0, 10]$  in steps of 0.1,  $A_V = 0$ –4 mag in steps of 0.1, and 51 age steps from 0 to the age of the Universe (see Bolzonella et al. 2000), the SFH of the decreasing models is sampled with  $\tau = (10, 30, 50, 70, 100, 300, 500, 700, 1000, 3000, \infty)$  Myr.

For all the above combinations we compute the  $\chi^2$  and the scaling factor of the template, which provides information about the SFR and stellar mass ( $M_{\star}$ ), from the fit to the observed SED. Minimisation over the entire parameter space yields the best-fit parameters. To determine both confidence intervals (68%) and medians for all the parameters, we ran 1000 Monte Carlo simulations for each object by perturbing the input broadband photometry assuming the photometric uncertainties are Gaussian. This procedure yields the one or two dimensions probability distribution function of the physical parameters of interest both for each source and for the ensemble of sources.



**Fig. 1.** Comparison between photometric redshift and spectroscopic redshifts with in green, black, red and blue the  $U$ ,  $B$ ,  $V$ ,  $i$ -dropout. Disagreements within 68% confidence limit for  $U$ ,  $B$ ,  $V$  and  $i$ -dropout affect 9 objects with very good/good spectroscopic redshift, 6 uncertain and 5 unreliable. Large error bars are due to sources with maxima probability at low and high redshift.

#### 3.2. Redshift selection

We have compared our photometric redshifts against objects with known spectroscopic redshift taken from literature (Vanzella et al. 2005, 2006, 2008; Mignoli et al. 2005; Szokoly et al. 2004; Le Fèvre et al. 2004; Doherty et al. 2005; Wolf et al. 2004; Daddi et al. 2005; Cristiani et al. 2000; Strolger et al. 2004; van der Wel et al. 2005; Roche et al. 2006; Ravikumar et al. 2007; Teplitz et al. 2007; Xu et al. 2007; Gronwall et al. 2007; Nilsson et al. 2007; Hathi et al. 2008; Finkelstein et al. 2008; Yang et al. 2008; Popesso et al. 2009). The result is shown in Figure 1, respectively for the  $U$ ,  $B$ ,  $V$  and  $i$ -dropout samples, for which 42, 72, 50 and 14 spectroscopic redshifts are available.

To estimate our photometric redshift performance, we compute the median  $\Delta z/(1+z_{\text{spec}})$ , with  $\Delta z$ , the difference between the median  $z_{\text{phot}}$  and  $z_{\text{spec}}$ . For each sample and model, we obtain values from  $-0.02$  to  $0.09$ , with no significant differences among models. This result show that we recover redshift with a good accuracy, consistent with typical values found in other studies (Wuyts et al. 2009; Hildebrandt et al. 2010).

Combining results of DEC, DEC+NEB+Ly $\alpha$  and DEC+NEB models, we find 20 objects among all samples, whose median photometric redshift is inconsistent with the spectroscopic redshift within the 68% confidence limit. The GOODS MUSIC catalog provides quality flags for the spectroscopic redshift; among the objects with inconsistent redshifts 6 are very good, 3 good, 6 uncertain and 5 unreliable spectroscopic redshifts, leading to an estimated 5–11% of outliers in our samples. We also obtain objects with large error bars, which is due to a double peaked redshift probability distribution function with maxima at low and high redshift.

To eliminate low redshift contaminants and to have the most reliable sample at each redshift we proceed with a conservative cut: for  $U$ ,  $B$ ,  $V$  and  $i$ -dropouts, we take a lower limit for the median photometric redshift of  $z > 2$ ,  $z > 3$ ,  $z > 4$  and  $z > 5$  respectively, as derived from for the DEC, DEC+NEB+Ly $\alpha$  and DEC+NEB models. We obtain respectively 389 ( $\sim 88\%$ ), 705 ( $\sim 82\%$ ), 199 ( $\sim 72\%$ ) and 60 ( $\sim 91\%$ ) objects. Similar criteria applied with REF models leads to larger sample (5 to 13%) and with RIS models to similar samples with a maximum variation of the size of the sample of 4%. We notice that with this final selection, there is a significant overlap between the  $U$  and  $B$  dropout with 96 objects in both samples.

Table 1 shows median redshifts and 68% confidence limits for REF, DEC and RIS model. Considering +NEB+Ly $\alpha$  and +NEB models induces median redshift variation not higher than 0.1.

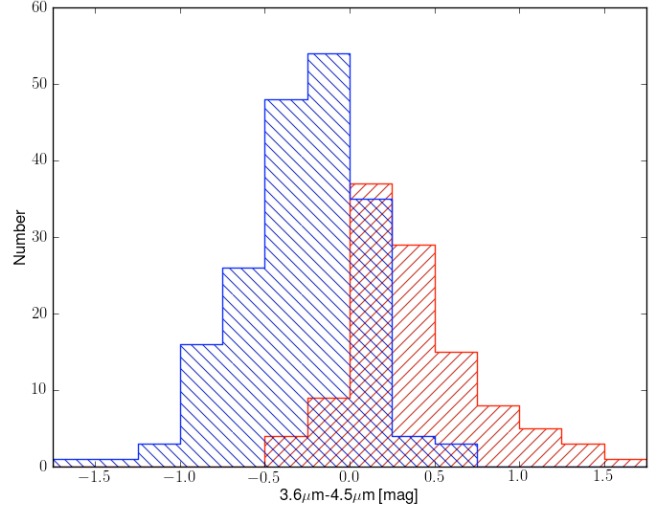
**Table 1.** Median redshift values and 68% confidence limits of the final samples.

	REF	DEC	RIS
$U$ -dropout	$3.32^{+0.25}_{-0.13}$	$3.33^{+0.26}_{-0.14}$	$3.30^{+0.25}_{-0.13}$
$B$ -dropout	$3.88^{+0.38}_{-0.41}$	$3.79^{+0.39}_{-0.44}$	$3.78^{+0.38}_{-0.43}$
$V$ -dropout	$4.94^{+0.31}_{-0.33}$	$4.81^{+0.39}_{-0.23}$	$4.81^{+0.39}_{-0.25}$
$i$ -dropout	$6.00^{+0.48}_{-0.29}$	$6.00^{+0.46}_{-0.32}$	$6.00^{+0.56}_{-0.33}$

## 4. Results

### 4.1. Two LBG categories revealed

As this is the first time large samples of LBGs are analysed with SED fits including the effects of nebular emission, we have examined if this leads to better fits and by how much. Our main result from this comparison is that typically 60–70% of the galaxies are better fit with nebular emission (option +NEB or +NEB+Ly $\alpha$ ) than without. This is found quite independently of the adopted SFH and for all samples, i.e. from  $z \sim 3$  to 6. In other words  $\sim 35\%$  of the objects are best fitted *without* taking account of nebular emission. This fraction is independent of properties such as the absolute UV magnitude  $M_{1500}$  or the number of filters available. Furthermore, all SF histories (REF, DEC, RIS model sets) yield approximately the same percentages (30%–39%). Finally, this is not only a statistical property,



**Fig. 2.**  $3.6\mu\text{m}$ – $4.5\mu\text{m}$  color histogram for a sub-sample of  $z \in [3.8, 5]$  objects. In blue, best fitted objects with nebular emission and red, best fitted objects without nebular emission, both for decreasing SF.

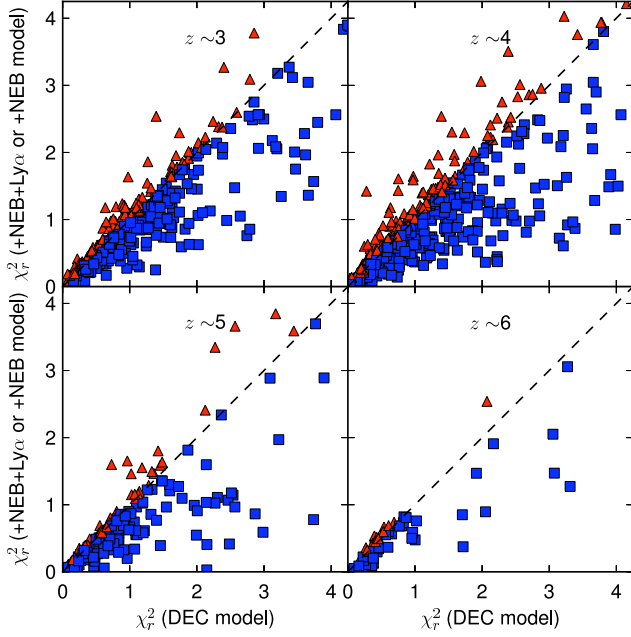
but the vast majority of objects can be assigned to such a category. Indeed, among the  $U$ ,  $B$ ,  $V$  and  $i$ -dropouts, we find 68%, 71%, 71% and 77% common objects best fitted without nebular emission and 85%, 80%, 94% and 88% common objects best fitted taking account nebular emission.

Since this distinction in two groups is fairly model- and redshift- independent there must be a physical explanation for it. The easiest and most natural explanation is found when considering a sub-sample of objects over a restricted redshift interval. Indeed, since H $\alpha$  is a strong line at 656.4 nm (rest-frame) and few strong lines are found longward of it, this line must affect the  $3.6$ – $4.5 \mu\text{m}$  color for objects between  $z=3.8$  and  $z=5$  (cf. Shim et al. 2011). We therefore selected  $B$ -dropout objects with available  $3.6\mu\text{m}$  and  $4.5\mu\text{m}$  data (excluding non-detections) and with median redshift between 3.8 and 5. We obtain a subsample of 303 objects, for which again  $\sim 35\%$  are better fit without nebular emission. This should thus be a representative subsample of all galaxies studied here. Figure 2 shows that the objects best fitted with nebular emission have a systematically bluer  $3.6\mu\text{m}$ – $4.5\mu\text{m}$  color than those better fit excluding nebular effects. This shows that objects better fit with models accounting for nebular emission do indeed show strong H $\alpha$  emission lines. This is not a trivial finding, since these models also allow for ages/SF histories where nebular emission is absent/insignificant. We therefore conclude that the objects best fit with models accounting for nebular emission ( $\sim 60$ – $70\%$ ) correspond to galaxies with “strong” emission lines, whereas the rest shows few or no discernible signs of emission lines.

It is interesting to note that both Yabe et al. (2009) and Shim et al. (2011) also found  $\sim 70\%$  of a sample of LBGs at  $z \sim 5$  and 4 respectively with a  $3.6 \mu\text{m}$  excess, which they explained by the presence of H $\alpha$  emission. While their samples include  $\sim 100$  (70) such galaxies, we have  $\sim 300$  galaxies at  $z \approx 3.8$ – $5$  with direct empirical evidence for strong H $\alpha$  emission. In addition our SED fits suggest that a similar percentage of objects with “strong” emission lines exists over the entire examined redshift range ( $z \sim 3$ – $6$ ). Below we will show that this interpretation is perfectly consistent with the differences found for the physi-



cal parameters of these galaxies, and we will propose a physical explanation for the existence of these two groups of LBGs.

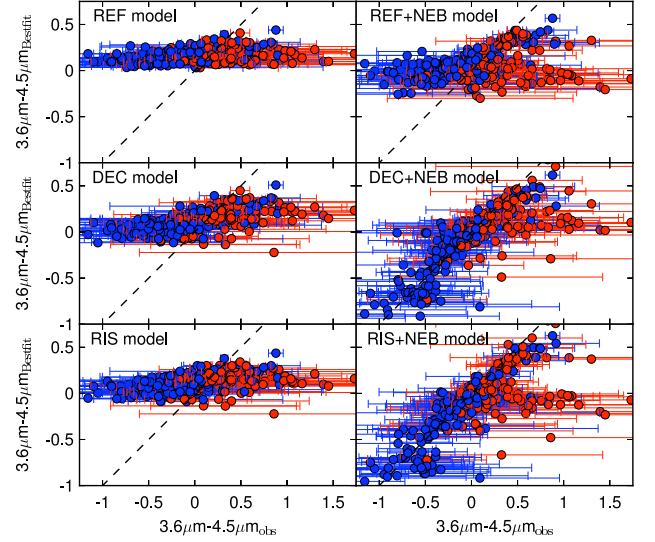


**Fig. 3.** Comparison between DEC model  $\chi_r^2$  and DEC+NEB+Ly $\alpha$ /+NEB model  $\chi_r^2$ . Red triangle: DEC model best fit and blue square: DEC+NEB+Ly $\alpha$ /+NEB best fit. Blue squares and red triangles underly the two LBGs categories, respectively “strong” nebular emitters and “weak” nebular emitters.

#### 4.2. Fit quality and constraints on the star formation histories

To compare the fit quality of the different models, we compare values of  $\chi_r^2$  ( $\chi^2$  value divided by the number of filters minus 1). At each redshift SEDs are systematically better fitted with RIS and DEC model sets (considering or not nebular emission) in comparison with REF model sets.  $\chi_r^2$  values are on average 20–40% lower for RIS model sets and 25–40% lower for DEC model sets, which also shows that DEC model sets fit slightly better than RIS model sets. As shown in Figure 3 for declining SF and all redshifts, “strong” nebular emitters show a large improvement in  $\chi_r^2$  when they are fitted with models considering nebular emission ( $\chi_r^2 \sim 30$  to 55% lower on average). At the opposite, “weak” nebular emitters show a slight improvement of their  $\chi_r^2$  when they are fitted without nebular emission (10 to 30% on average). Models with nebular emission are able to provide significantly better fits for “strong” nebular emitters and a more or less similar fit quality for “weak” nebular emitters.

For the subsample of objects with redshifts  $z \sim 3.8$ –5 discussed above, the observed  $3.6\mu\text{m}$ – $4.5\mu\text{m}$  provides also an interesting constraint on star formation history and nebular emission. Indeed, as shown in Fig. 4, models without nebular emission are unable to reproduce the range of observed  $3.6\mu\text{m}$ – $4.5\mu\text{m}$  colors at  $z \sim 4$ . The REF+NEB model (constant star formation) is also unable to reproduce observations while the DEC+NEB and RIS+NEB models provide fits fully consistent with this color (+NEB+Ly $\alpha$  option leads to similar results), with DEC+NEB providing even better results than RIS+NEB. Abandoning the age limitation (age  $> 50$  Myr) for the REF+NEB model



**Fig. 4.** Comparison between observed  $3.6\mu\text{m}$ – $4.5\mu\text{m}$  color and best fit color for a sub-sample of  $z \in [3.8, 5]$  objects for SFHs with and without nebular emission.

would allow this model to provide fits consistent with observations. We need to have very large H $\alpha$  equivalent widths (EWs) to reproduce the color of “strong” nebular emitters. This can be obtained for any SFH with young ages (median of  $\sim 20$  Myr for DEC+NEB and RIS+NEB). On the other hand for “weak” nebular emitters small EW(H $\alpha$ ) are necessary to reproduce their red color. Since both for rising and constant SFs no physical process is able to suppress effectively nebular emission, DEC+NEB(+Ly $\alpha$ ) is the model which best fits both LBG categories. However, considering errors in color estimation, RIS+NEB(Ly $\alpha$ ) and REF+NEB(+Ly $\alpha$ ) provide also acceptable fits for “weak” nebular emitters. Further constraints and results on the timescales of the exponentially declining star formation histories are discussed in Sect. 4.3.5.

Finally, among objects better fitted with nebular emission we have found a shift between those best fit with or without Ly $\alpha$  (i.e. between +NEB and +NEB+Ly $\alpha$  option) with redshift, in the sense that higher- $z$  galaxies favor a larger fraction of objects with Ly $\alpha$  emission. This shows that SED fitting is also sensitive to Ly $\alpha$  emission, a finding we have demonstrated and discussed in detail in Schaerer et al. (2011).

#### 4.3. Physical properties

We now turn to discuss the main physical properties (stellar mass, SFR, age, attenuation, plus the star formation timescale where appropriate) of the LBGs and their dependence on model assumptions. The median values and uncertainties of the physical parameters derived in several bins of UV magnitude  $M_{1500}$  for all our samples, and using all nine combinations of model assumptions, are listed in Tables A.1, A.2, and A.3. For each physical parameter we will now describe the median properties and their model dependence, explain their origin, and compare the behaviour of the individual values. Furthermore we will examine possible correlations between derived and observed parameters. To do this we here chose the largest subsample, consisting of 705  $B$ -drop ( $z \sim 4$ ) galaxies, since overall the same trends/differences are found at all redshifts (except stated otherwise). In Sect. 4.4 we then discuss the redshift evolution of the physical parameters and their model dependence.

#### 4.3.1. Absolute UV magnitude

In what follows the absolute UV magnitude  $M_{1500}$  refers to the absolute magnitude at 1500 Å. To determine it for each object, we use the integrated SED flux in an artificial filter of 200 Å width centered on 1500 Å. Using the  $V$ -band magnitude for  $U$ -drop,  $i$ -band for  $B$ -drop, and  $z$ -band for  $V$  and  $i$ -drop samples and spectroscopic redshift when available, to estimate the UV magnitude (Stark et al. 2009) leads to no significant difference on  $M_{1500}$ , except for one  $B$ -drop, two  $V$ -drop and one  $i$ -drop galaxy. These are in fact objects with spectroscopic redshift identification at low redshift which pass our selection, and as they represent less than 1% (slightly more for  $i$ -drop) of each sample, we consider that they can not alter significantly our conclusions. In passing we note that the number of objects in each UV magnitude bin listed in Tables A.1, A.2 and A.3 can somewhat change from one model to another, mostly due to small differences in photometric redshifts.

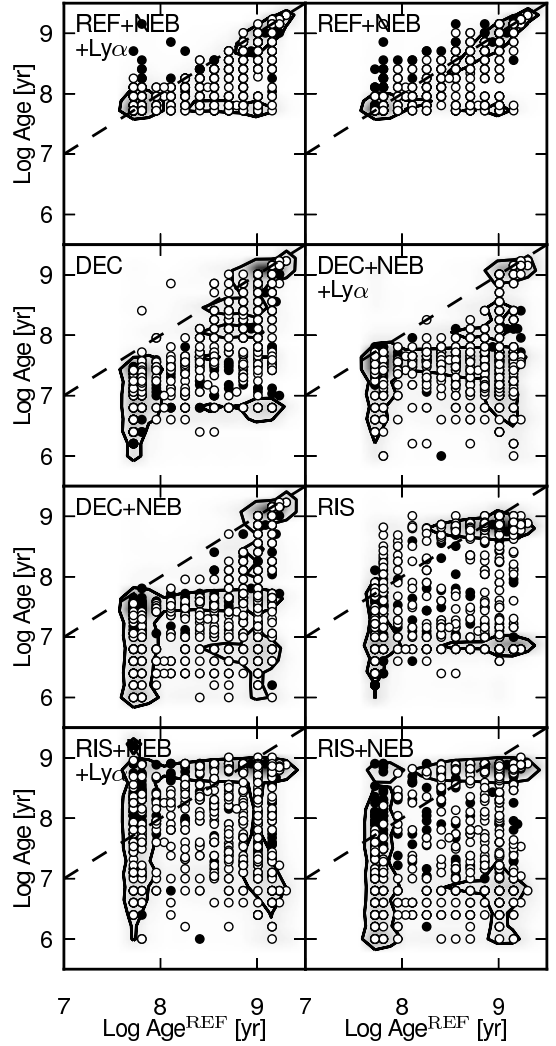
#### 4.3.2. Age

Overall, the ages of individual  $z \sim 4$  LBGs derived from the different models span a wide range, typically from  $\sim 4$  Myr (if no lower limit is specified) to  $\sim 1$  Gyr, the maximum age at this redshift (see Fig. 5). A wide age range is found for all 9 model sets.

The individual ages and the resulting median age of the sample depend quite strongly on the model assumptions. As can be seen in Tables A.1–A.3, the assumption of declining (DEC) or rising (RIS) star formation histories lead – for models without nebular emission – to median ages younger by a factor 5–10 compared to constant star formation (REF). These differences are much larger than the typical age uncertainty ( $\sim 0.15$  dex) found for the REF model. The reason is that with a constant SF, galaxies keep a high UV rest-frame flux and if the observed SED shows the presence of a Balmer break, older ages are necessary to have a sufficient population of evolved stars to reproduce the observed break. For a declining or rising SF, much younger ages are sufficient to obtain a suitable evolved population. Indeed the observed median ratio of the optical/UV flux is high and consistent with a Balmer break at each redshift (see González et al. 2010; Lee et al. 2011).

As already shown in Schaerer & de Barros (2009, 2010), emission lines (mostly [O III]  $\lambda\lambda 4959, 5007$  plus H $\beta$ ) can mimic a Balmer break for a constant or decreasing SF, and our results at  $z \sim 3$ –5 demonstrate this effect for a large sample: the inclusion of nebular emission (with or without Ly $\alpha$ ) leads to systematically younger median ages for the REF and DEC model sets (e.g. by a factor 2–5 for the REF model; see Fig. 5). This is not the case for the rising star formation history: RIS+NEB+Ly $\alpha$  and RIS+NEB models lead to older or similar ages, for reasons explained below. Furthermore, as shown in Schaerer & de Barros (2010), taking into account nebular emission increases the uncertainty on the age, especially for REF and RIS model sets, from 0.15 to 0.27 dex and from 0.39 to 0.55 dex respectively ( $\sim 0.4$  dex for DEC model sets). However, these uncertainties are sufficiently small to distinguish differences in age among the three SFHs (constant, decreasing and rising) and between models with/without nebular emission.

To discuss the effect of nebular emission on age it is useful to consider the two LBG categories, “weak” and “strong” line emitters, separately. Median ages of the “weak” nebular emitters increase systematically when we take into account nebular emission, for all considered star formation histories. This is nat-



**Fig. 5.** Composite probability distribution of age for REF model and age for all other models at  $z \sim 4$ . The points overlaid show the median value properties for each galaxy in the sample, black dots for “weak” nebular emitters and white dots for “strong” nebular emitters. The overlaid contour indicates the 68% integrated probabilities on the ensemble properties measured from the centroid of the distribution.

ural, since in all cases the equivalent widths of the optical emission lines decrease with time. Hence increasing ages minimizes their contribution. For the RIS model, the ages of “weak” nebular emitters are much more increased than for REF and DEC models, by a factor 10–20 at  $z \sim 4$  (compared to 2–5 times for REF and DEC models). This strong increase outweighs the age decrease of the “strong” emitters, which explains why the median age of the whole population increases or does not change, when we take into account nebular emission for the rising star formation history.

Quite generally, and for all SFHs, we find for “strong” nebular emitters median ages, which are systematically younger when nebular emission is included, because of a stronger effect already discussed in Schaerer & de Barros (2009, 2010) and confirmed here: strong lines mimic a Balmer break. Only at  $z \sim 6$  and for DEC+NEB and RIS+NEB models do we find older ages. Balmer break seems to be absent in these galaxies, which are fit

with a strong UV flux (models with no nebular emission) and/or strong Ly $\alpha$  emission (+NEB+Ly $\alpha$ ). For DEC or RIS models, this implies a young population, and for DEC/RIS+NEB+Ly $\alpha$  too. For DEC/RIS+NEB models, the lack of Ly $\alpha$  emission must be balanced by an increase of UV flux, which implies a younger stellar population, but a younger population leads to strong lines which mimic a strong Balmer break, this is why age cannot be as young than for +NEB+Ly $\alpha$  models.

In absolute terms, the derived ages of “weak” and “strong” nebular emitters are fairly similar when constant SF is assumed. For decreasing and rising SF, the median ages of “strong” nebular emitters are systematically older than those of the “weak” emitters for models without nebular emission, while the opposite is found for models with nebular emission. In the former case (no nebular emission) the “strong” emitters are fitted with a strong Balmer break, in the latter case with strong lines. In any case, the age differences seem to confirm an intrinsic difference between these two LBG categories.

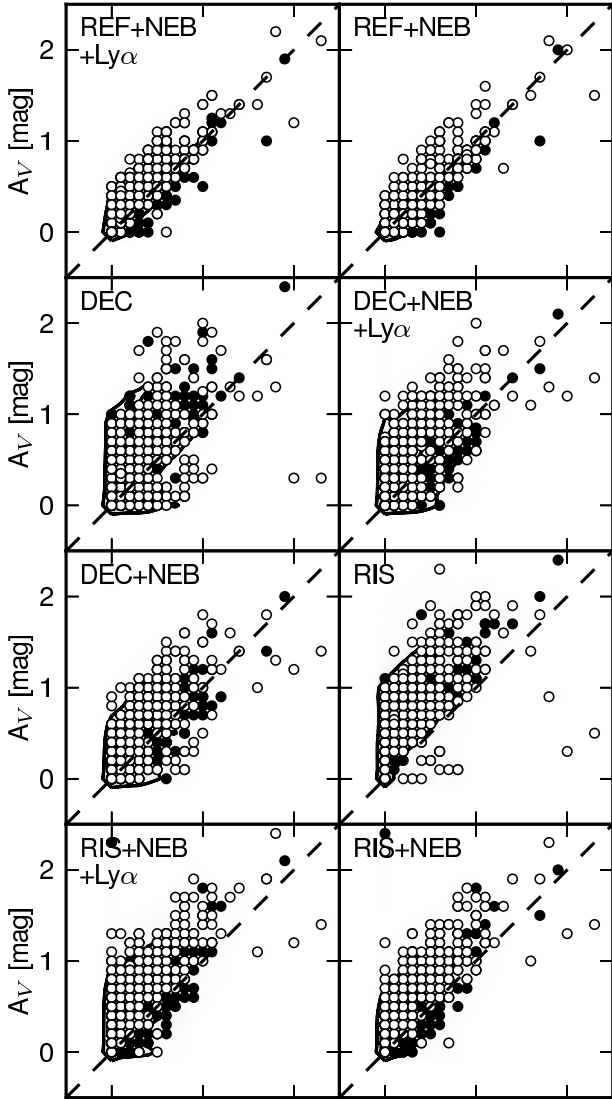


Fig. 6. Same as Figure 5 for  $A_V$ .

#### 4.3.3. Reddening

The attenuation of individual  $z \sim 4$  LBGs derived from the SED fits ranges from  $A_V = 0$  to a maximum of  $\sim 1.5$ – $2$  mag for few objects, as shown in Fig. 6. Although all model sets yield a similar range of attenuation, relatively large systematic differences, which we will just discuss, are found between them.

Considering or not nebular emission for a given SFH lead to variation no larger than 0.2 mag. Comparing constant SF with other star formation histories, the median  $A_V$  is higher for declining SF (+0.2 mag) and for rising SF (+0.5 mag), which can be partially explained by the well-known degeneracy between reddening and age, since declining and rising SF lead to younger median ages than constant SF (cf. above). Furthermore, since young stars always dominate the UV flux for rising star formation histories, a higher attenuation is required to fit the observations, compared to models with constant SF, as already pointed out by Schaefer & Pelló (2005). Taking into account nebular emission leads on average to variations no larger than 0.1 mag in the reddening for constant SF; a slight increase is found when the Ly $\alpha$  line is included (DEC+NEB+Ly $\alpha$  model, assuming the maximum case B intensity for Ly $\alpha$ ), which is explained by the additional flux from the Ly $\alpha$  line.

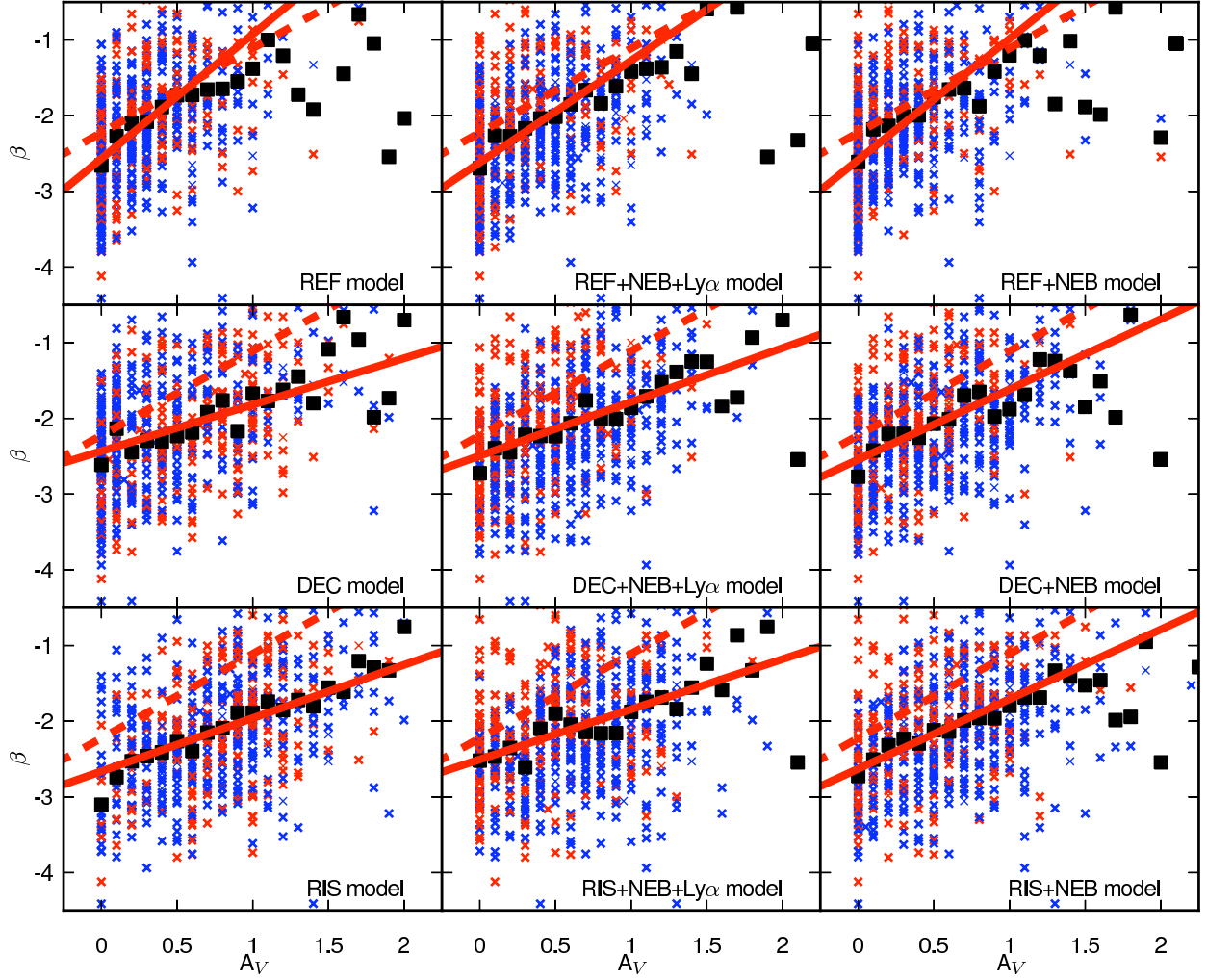
On the other hand, for rising star formation rate, taking into account nebular emission leads to a systematically lower median  $A_V$ , since the contribution of strong nebular lines in the optical and infrared (rest-frame) leads to redder SEDs. These results have to be taken with some caution since typical errors on  $A_V$  for individual objects are 0.1 mag for constant SF and  $\sim 0.2$  mag for decreasing and rising SF, which does not allow a clear distinction for example between REF and DEC models.

The inclusion of nebular emission with constant star formation history does not lead to any significant change in dust reddening for “weak” nebular emitters. On the other hand, for declining and rising SF, the extinction decreases strongly (by  $\sim -0.2$  to  $-0.5$  mag in  $A_V$ ) when we consider nebular emission. Since these objects seem to have intrinsically no discernible signs of emission lines, models including nebular emission fit by minimizing equivalent widths, and so, by minimizing SFR and UV flux. Model set based on constant star formation history (REF/REF+NEB/REF+NEB+Ly $\alpha$  models) do not allow sufficient variations to modify the reddening estimation. Median  $A_V$  between  $z \sim 3$  and  $\sim 5$  for “strong” nebular emitters increases when we consider decreasing star formation (+0.2 to +0.5 mag), while it decreases for rising SF ( $-0.1$  to  $-0.2$  mag). Indeed, fitting “strong” nebular emitters with DEC model leads to increase the purely stellar continuum contribution in the optical/IR and therefore increase the age and decrease SFR and UV flux. Rising SF leads to very large equivalent widths (in comparison with other SFHs) in optical and near-infrared (rest-frame), which redden SEDs and allow lower values of  $A_V$ .

At  $z \sim 6$ , effects of modeling with nebular emission are different: for decreasing and rising model sets, considering nebular emission leads to a decrease of median  $A_V$  for both “weak” and “strong” nebular emitters, respectively with  $\sim -0.6$  and  $\sim -0.2$  mag. SED fits with nebular emission lead to an important contribution of nebular lines longward UV, and strong lines being associated with strong UV flux, an additional amount of dust attenuation is required.

#### 4.3.4. Reddening and UV-slope

Since often the observed UV slope  $\beta$  is used to measure the attenuation in LBGs, it is interesting to examine how the attenuation



**Fig. 7.** UV-continuum slope  $\beta$  vs  $A_V$  for all models at  $z \sim 4$ . Each cross shows the median value properties for each galaxy, blue for “strong” nebular emitters and red for “weak” nebular emitters. Black squares are median values by bins of  $0.1 A_V$  mag. The red dashed line is the relation between extinction (Calzetti et al. 2000) to a given  $\beta$  if the base spectrum is a young star-forming galaxy of constant star-formation (Bouwens et al. 2009). The red line is a linear fit among the whole composite probability distribution function.

derived from the SED fits based on the different model assumptions compares with  $\beta$ . Such a comparison is shown in Fig. 7 for the nine model sets applied to the  $B$ -drop sample. The UV slope has been determined using the same filters and relations as Bouwens et al. (2009). Figure 7 shows that there is a significant trend of increasing  $\beta$  with  $A_V$ , as expected, albeit with a large scatter. We have done linear fits to the 2D composite probability distribution function, which yields the mean relations indicated in the plot by red lines. For comparison the “standard” relation between  $\beta$  and  $A_V$ , here taken from Bouwens et al. (2009), is also shown. As expected, our relations agree well with the “standard” one for models assuming constant star formation and ages  $> 50$  Myr (REF model sets), since this corresponds to the main assumption made to derive the standard  $\beta$ –reddening relation. For a given SF history, the differences between the three options with/without nebular emission can be explained by the behaviour of  $A_V$  discussed above. Since all models with declining and rising star formation histories yield somewhat higher reddening on average (cf. above), a relation steeper than the standard one is found. Since the obtained relations are fairly similar

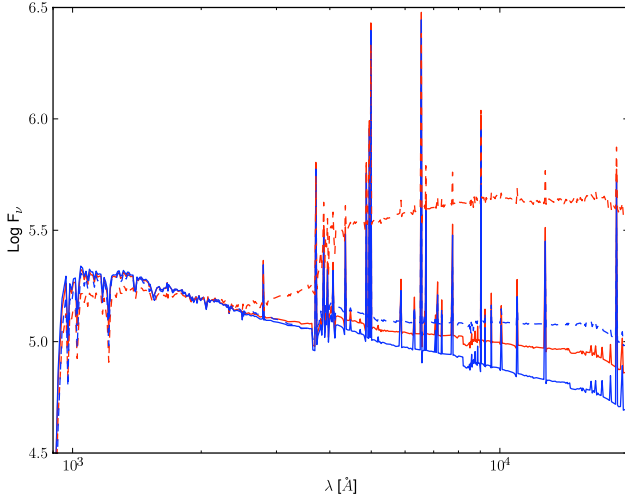
we can combine them to obtain the following mean relations between  $\beta$  and  $A_V$  for the three cases, (1) modeling without nebular emission, (2) +NEB+Ly $\alpha$  and (3) +NEB:

$$A_V = 1.54 \times (\beta + 2.54) \quad (1)$$

$$A_V = 1.47 \times (\beta + 2.49) \quad (2)$$

$$A_V = 1.09 \times (\beta + 2.58) \quad (3)$$

The last relation (Eq. 3) is probably the most appropriate one, since it combines the models which best fit the data (i.e. models including nebular emission but no strong Ly $\alpha$  for the majority of the galaxies, cf. Schaerer et al. 2011). It should be reminded that these relations assume a Calzetti attenuation law; to translate this into the color excess one has  $E(B - V) = A_V/R_V$ , where  $R_V = 4.05$ . In short, from our new relations derived from a subsample of 705  $B$ -drop galaxies using various star formation histories we find that LBGs with a given UV slope have higher attenuation than derived from the commonly used  $\beta$ – $A_V$  relation. For typical average UV slopes of  $\beta \sim -2.2$  ( $-1.7$ ) found for faint (bright)  $z \sim$



**Fig. 8.** Theoretical SEDs in the rest-frame, normalized at  $2000\text{\AA}$ , for models with  $\text{EW}(\text{H}\alpha) = 500\text{\AA}$  (solid lines) and  $100\text{\AA}$  (dashed lines) and different star formation timescales (constant SF in red and  $\tau = 10\text{ Myr}$  in blue). For the models with  $\text{EW}(\text{H}\alpha) = 500\text{\AA}$  the ages are 52 Myr and 16 Myr respectively, for  $\text{EW}(\text{H}\alpha) = 100\text{\AA}$ , 2.1 Gyr and 35 Myr.

4 galaxies, this translates to an increase of the UV attenuation by a factor  $\sim 3$ .

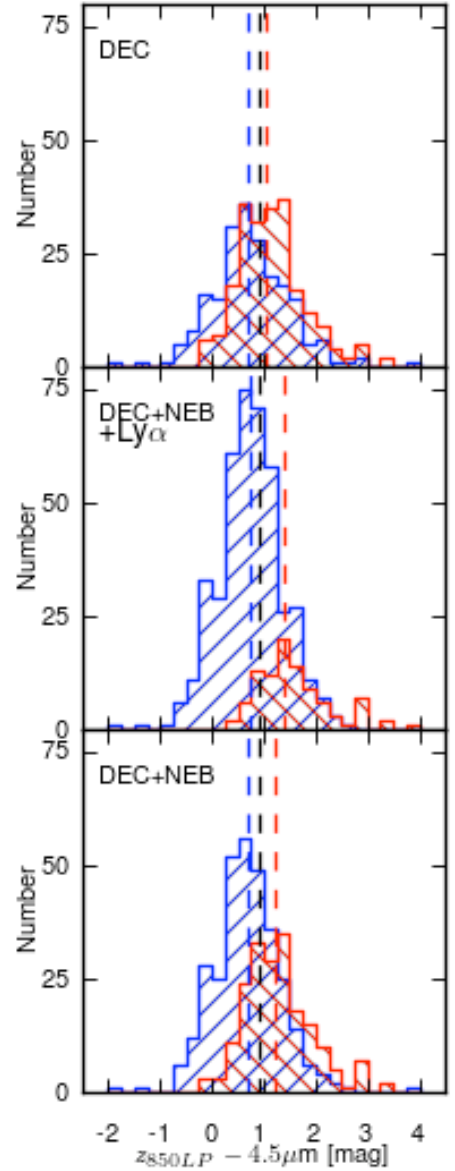
#### 4.3.5. Star formation timescale

For the models with exponentially declining star formation histories (DEC models), which are found to provide the best fits for the majority of objects – although not always by large margins – it is of interest to examine the resulting timescales  $\tau$  and the uncertainties on this quantity. As a reminder, the DEC model set considers 10 different star formation timescales  $\tau \in [10, 3000]\text{ Myr}$  plus the limiting case of  $\tau = \infty$  corresponding to constant star formation.

For all options with/without nebular emission (DEC, DEC+NEB+Ly $\alpha$  and DEC+NEB) we find median values of  $\tau$  between 10 and 300 Myr in the different UV magnitude bins (cf. Table A.2). Models including nebular emission favour on average somewhat shorter timescales than those without.

Although the timescales found are relatively short compared to the dynamical timescale  $t_{\text{dyn}} \sim 40\text{ Myr}$  at  $z \sim 4$  – the uncertainties on  $\tau$  are quite large.

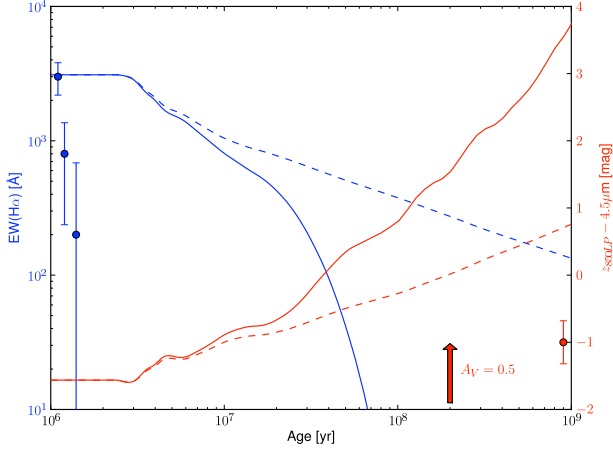
What constrains the SF timescales and why are short timescales preferred? Since  $\text{SFR} \propto \exp(-t/\tau)$  two (or more) observational constraints are needed to determine both the age  $t$  and timescale  $\tau$ . Let us first consider the case of models including nebular emission and examine galaxies with  $z \sim 3.8\text{--}5$ . In this case we find that  $t$  and  $\tau$  are mostly constrained by a combination of the  $(3.6\text{--}4.5)\text{ }\mu\text{m}$  color tracing  $\text{EW}(\text{H}\alpha)$ , and by a UV/optical (rest-frame) color. This works as follows. As already discussed above, a measurement of the  $3.6\text{--}4.5\text{ }\mu\text{m}$  color for galaxies at  $z \sim 3.8\text{--}5$  reflects the  $\text{H}\alpha$  equivalent width. However, it is well known that a given equivalent width can be obtained with different values of  $t/\tau$  (cf. Fig. 10). To illustrate this we show in Fig. 8 the predicted SEDs for galaxies with the same  $\text{EW}(\text{H}\alpha) = 100\text{ (500) }\text{\AA}$  but different SF timescales ( $\tau = 10\text{ Myr}$  and  $\infty$ ). It is obvious that the main feature allowing to lift this degeneracy is the ratio of the UV/optical flux. At  $z \sim 4$  this ratio is e.g. reflected



**Fig. 9.**  $z_{850LP} - 4.5\text{ }\mu\text{m}$  observed color histogram at  $z \sim 4$ . In red, best fitted objects with  $\tau \geq 100\text{ Myr}$  and blue, best fitted objects with  $\tau = 10\text{ Myr}$ , for the three models with decreasing SF. The blue, red and black dashed lines show respectively median color for objects best fitted with  $\tau = 10\text{ Myr}$ , with  $\tau \geq 100\text{ Myr}$  and for the whole sample. For the two subsamples, KS test for the three models gives  $p = 7.8 \cdot 10^{-8}$ ,  $p = 2.6 \cdot 10^{-17}$  and  $p = 1.3 \cdot 10^{-16}$  respectively for DEC, DEC+NEB+Ly $\alpha$  and DEC+NEB models, showing that the two subsamples are not drawn from the same population.

by the  $z_{850LP} - 4.5\text{ }\mu\text{m}$  color, whose evolution with  $t$  and  $\tau$  is also shown in Fig. 10. Therefore these two colors will provide good constraints on  $t$  and  $\tau$ , provided they are not strongly affected by reddening (cf. below). Indeed within typical 68% error bars, these two extreme SFHs can be discriminated by their color and  $\text{EW}(\text{H}\alpha)$  for  $t \gtrsim 20\text{--}30\text{ Myr}$ . For younger ages, the uncertainties in both  $\text{EW}(\text{H}\alpha)$  and color do not allow a clear separation. A posteriori we can verify that the objects best fitted with “long” timescales do indeed statistically differ from those with “short” timescales. Indeed Fig. 9 shows a statistically significant differ-



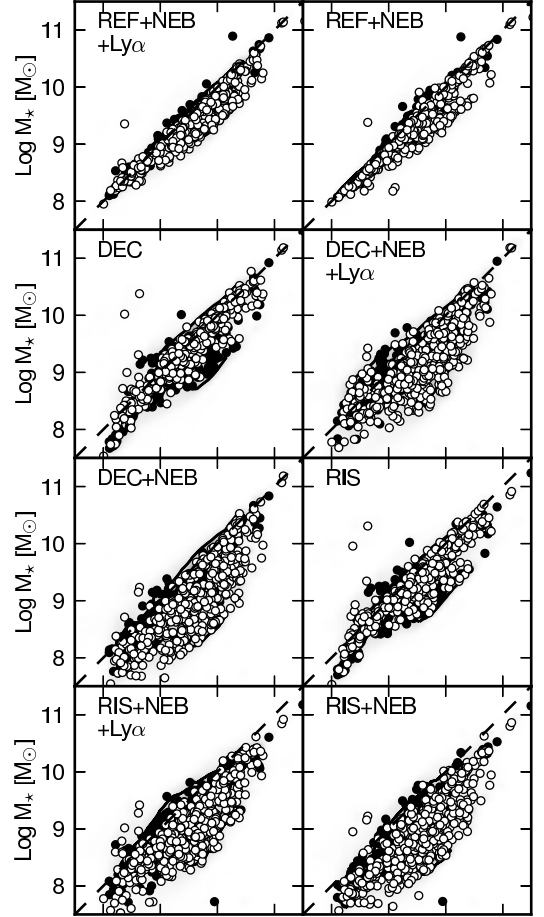


**Fig. 10.** Evolution of  $\text{EW}(\text{H}\alpha)$  (blue) and  $z_{850\text{LP}} - 4.5\mu\text{m}$  color (red) with age for  $\tau = 10$  Myr (solid lines) and for  $\tau = \infty$  (dashed lines). Typical error bars, shown on the left for  $\text{EW}(\text{H}\alpha)$  and on the right for the  $z_{850\text{LP}} - 4.5\mu\text{m}$  color, have been derived from error estimation on measured fluxes at  $z \sim 4$ . The effect of reddening ( $A_V = 0.5$ ) on  $z_{850\text{LP}} - 4.5\mu\text{m}$  is shown with the red arrow.

ence between galaxies best fitted with  $\tau = 10$  Myr, which are bluer in  $z_{850\text{LP}} - 4.5\mu\text{m}$ , and  $\tau > 100$  Myr showing redder colors.

We also have to consider dust attenuation, which can increase  $z_{850\text{LP}} - 4.5\mu\text{m}$  as shown in Fig. 10 and introduce a degeneracy between SFHs. However, considering the observed median color and dust attenuation for the REF and DEC model sets, we would expect to have higher extinction for REF model sets than for DEC model sets, but we find the opposite with median  $A_V$  larger by 0.1–0.2 mag for DEC models. Furthermore, Figure 10 shows that it is more difficult to reproduce red color with a constant SF than with a declining SF. This shows that dust attenuation seems to be only weakly correlated with  $z_{850\text{LP}} - 4.5\mu\text{m}$  color at  $z \sim 4$ , allowing us to conclude that the ratio of UV/near-IR flux and equivalent widths of different emission lines (mainly  $\text{Ly}\alpha$ , OIII and  $\text{H}\alpha$ ) drive the choice of  $\tau$  for models with declining star formation. Despite the possibility to discriminate two extreme SFHs like constant SF and decreasing SF with  $\tau = 10$  Myr, large uncertainties on  $\tau$  estimates are found (from 0.7 dex for DEC+NEB+ $\text{Ly}\alpha$  model to  $\sim 2$  dex for DEC+NEB model), preventing us from deriving strong constraint on  $\tau$ . These uncertainties come mainly from reddening; indeed fixing reddening to an arbitrary value ( $A_V = 0$ ) leads also to a low median  $\tau$  ( $< 300$  Myr) but with lower typical uncertainties, from  $\sim 0.4$  dex for the DEC+NEB+ $\text{Ly}\alpha$  model to  $\sim 0.6$  dex for the DEC model.

Interestingly, our models with declining SF histories indicate a possible increase of the timescale  $\tau$  with UV luminosity, and hence also with stellar mass, but only for “strong” nebular emitters. It is tempting to suggest that this could be due to a decrease of the feedback efficiency with increasing galaxy mass, since the star formation timescale is likely related to the dynamical one and modulated by feedback (e.g. Wyithe & Loeb 2011). We do not see any evolution of the timescale for the “weak” nebular emitters, which is easily explained by weaker constraints due to the absence of strong distinctive features. This explains why the trend of  $\tau$  with UV magnitude cannot be seen in Table A.2, where the combined data for the entire sample is listed.

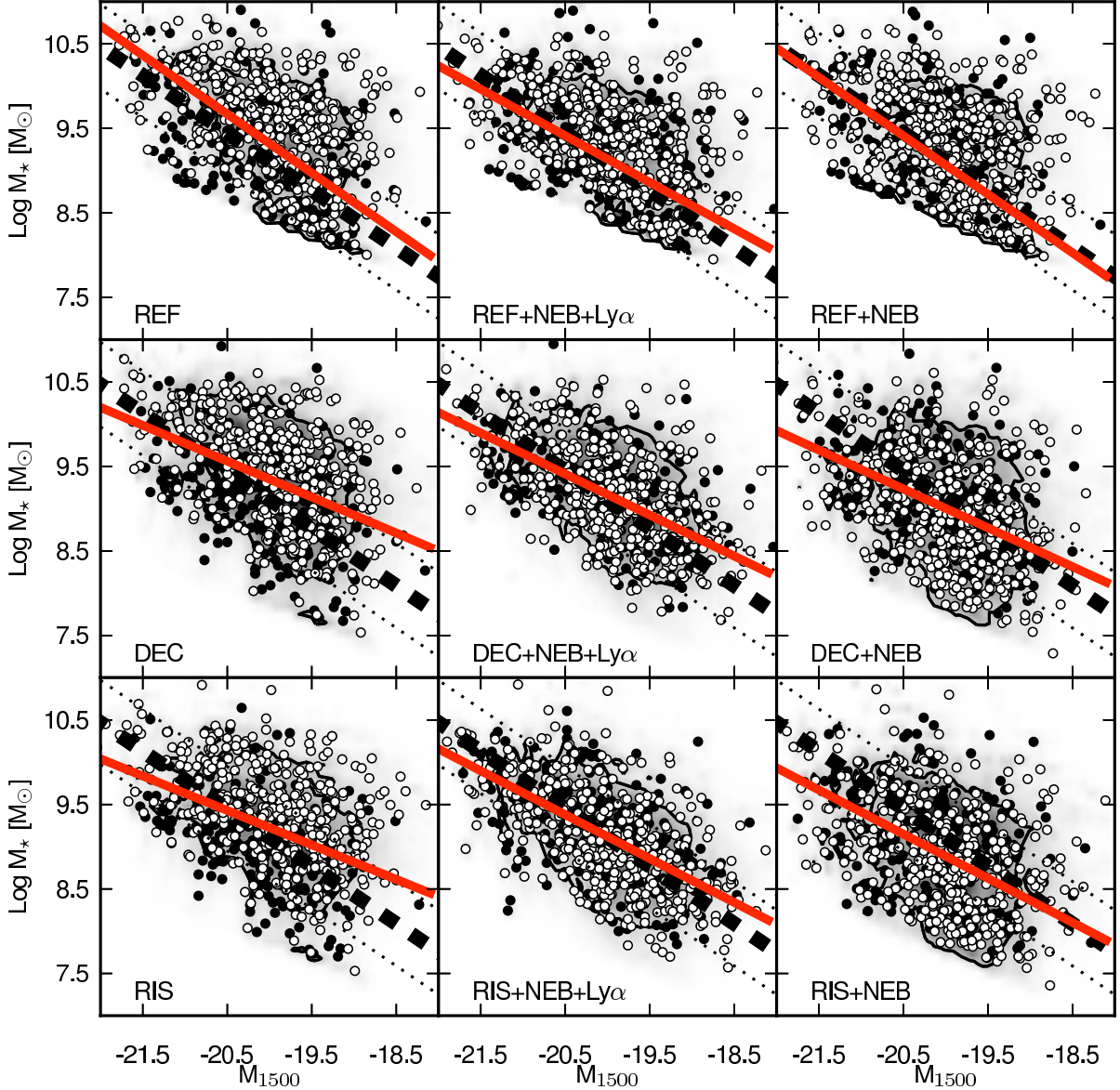


**Fig. 11.** Same as Figure 5 for  $M_*$ .

#### 4.3.6. Stellar mass

Stellar mass is generally considered the most reliable parameter estimated by SED fitting, since relatively small differences are found when varying assumptions like the star formation history or dust extinction. Finlator et al. (2007) estimates differences due to different assumptions on SFHs typically not higher than 0.3 dex and Yabe et al. (2009), adding effects of metallicities and extinction laws, estimates differences to be not higher than  $\sim 0.6$  dex. Our comparison of stellar masses based on the different model assumption is shown in Fig. 11 and in Tables A.1–A.3. Our results confirm earlier results on the dependence of stellar mass with the assumed SFH. Indeed, from  $z \sim 3$  to  $z \sim 6$  the median stellar masses do not differ by more than  $\sim 0.3$  dex between different star formation histories, even with metallicity included as a free parameter. With respect to models for constant star formation and without nebular emission (REF) all other models and options (+NEB or +NEB+ $\text{Ly}\alpha$ ) lead systematically to lower stellar masses, with differences exceeding the typical uncertainties of  $\sim 0.15$  dex found for the REF model. Taking into account nebular emission for constant SF leads to differences of same order as the uncertainties, not larger than  $\sim 0.2$  dex. For models with decreasing or rising SF and nebular emission included, we obtain stellar masses lower by  $\sim 0.4$  dex on average compared to the model with constant star formation.

As could be expected, the masses of galaxies from the category of “strong” nebular emitters vary in the same manner, but more strongly than those of the “weak” emitters, when different

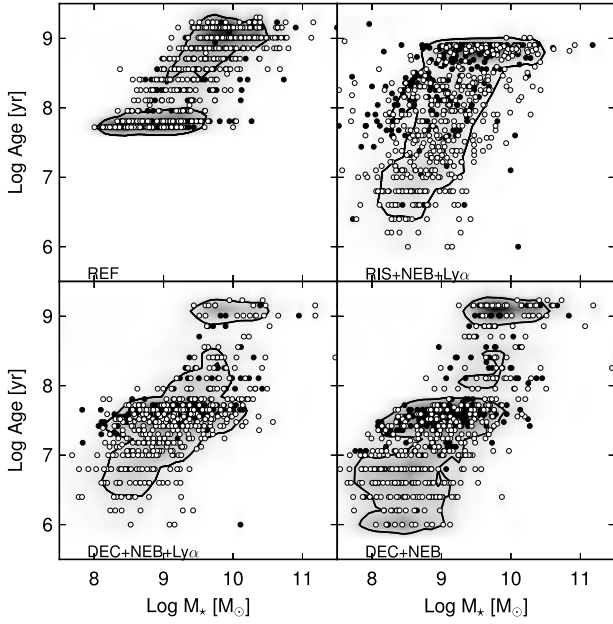


**Fig. 12.** Composite probability distribution of  $M_{1500}$  and  $M_*$  for all the models at  $z \sim 4$ . The black dashed line represents the  $M_*$ – $M_{1500}$  trend found by González et al. (2011) and black dotted lines shows a scatter of  $\pm 0.5$  dex. The solid red line shows a linear fit established taking into account the whole composite probability distribution. The points overlaid show the median value properties for each object in the sample, black dots for “weak” nebular emitters and white dots for “strong” nebular emitters. The overlaid contour indicates the 68% integrated probabilities on the ensemble properties measured from the centroid of the distribution.

SF histories and options on the treatment of nebular emission are taken into account. Typically masses decrease by  $\sim 0.4$ – $0.9$  ( $0.2$ ) dex for “strong” (“weak”) nebular emitters between constant SF and the other models. With declining and rising SF histories and SEDs including nebular emission, the categories of “strong” and “weak” nebular emitters show distinct median masses — the former being less massive than the latter — with differences up to  $\sim 0.6$  dex at  $z \sim 6$ . In contrast, models with constant SF and nebular emission do not yield significantly different masses between these LBG categories. This is due to the more restricted “dynamical range” allowed by these models (see e.g. Fig. 10).

In Figure 12, we show the stellar mass– $M_{1500}$  relation found for all our models at  $z \sim 4$ . For constant star formation with or without nebular emission we find, as expected, a relation in

good agreement with that established by González et al. (2011), within a scatter of  $\pm 0.5$  dex. Indeed, our REF model is based on assumptions similar to those of González et al. (2011), except for the metallicity (we assume  $Z = Z_\odot$  when they assume  $Z = 0.2Z_\odot$ ) and for the minimum age (we assume 50 Myr when they assume 10 Myr). Solar metallicity leading to  $\sim 0.06$  dex of increase in mass in comparison with  $0.2 Z_\odot$  and a higher minimal age leading to higher lower bound in  $M_*$ , the slight offset of our stellar mass– $M_{1500\text{\AA}}$  relation is easily explained. Overall, although differences exist in the  $M_*$ – $M_{1500}$  relation obtained between different model sets, our mean relation remains (within the scatter of  $\pm 0.5$  dex) fairly similar to the relation derived by González et al. (2011), regardless of our model assumptions.

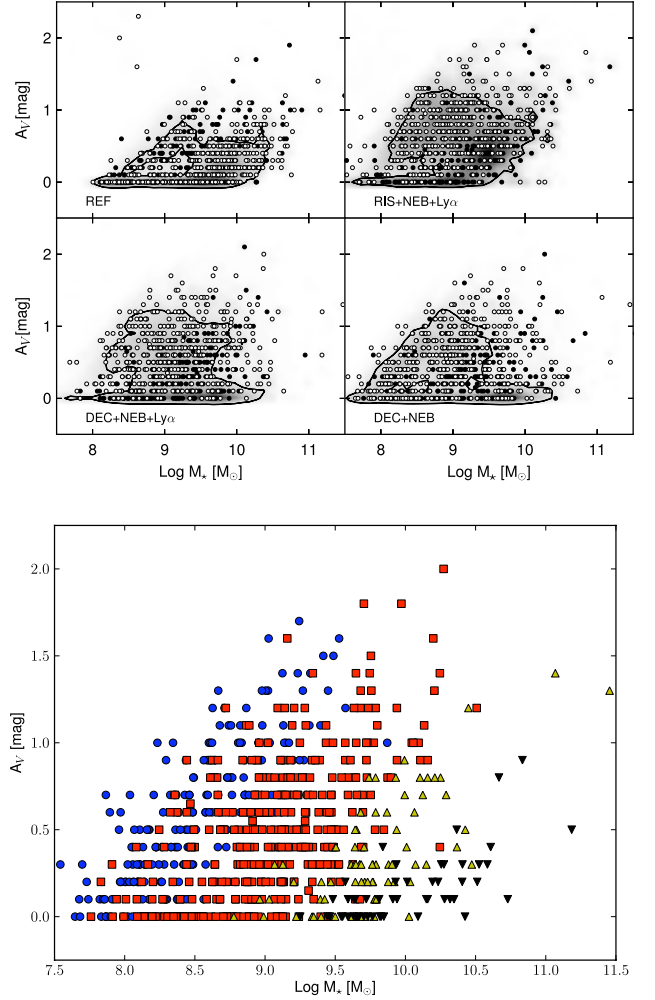


**Fig. 13.** Composite probability distribution of  $M_*$  and age for some models at  $z \sim 4$ . The points overlaid show the median value properties for each object in the sample, black dots for “weak” nebular emitters and white dots for “strong” nebular emitters. The overlaid contour indicates the 68% integrated probabilities on the ensemble properties measured from the centroid of the distribution.

A correlation between stellar mass and age is found for all the models, as shown in Figure 13. Since age estimation depends of the mass to light ratio, this relation is trivial if we describe star formation history with a monotonic function. Indeed, luminosity is fixed by both measured fluxes and redshift, and NIR data putting strong constrains on stellar mass estimation, the stellar mass–age relation simply reflects the increase of mass to light ratio with age. As previously explained, different assumption on the SFH lead to different trends between “weak” and “strong” nebular emitters, when analysed with SEDs including nebular emission: for variable star formation histories (both rising or declining) “weak” nebular emitters are found to be older and more massive on average than “strong” nebular emitters. In contrast, when constant star formation is assumed, the physical properties of the two populations do not differ.

In Figure 14 (top panel), we show the relation between the dust attenuation  $A_V$  and stellar mass, for a selected model set. For all models a similar trend is found with the median  $A_V$  increasing with galaxy mass, and a wide range of attenuations allowed between zero and  $\sim 1.5$  mag.

The bottom panel of Figure 14 allows a better understanding of the correlation between stellar mass and dust reddening, since it shows that the dispersion comes mainly from the age scatter. Indeed, for a given range of age, we find a clear trend of increasing extinction with increasing stellar mass. This trend has already been highlighted at lower redshift (eg. Buat et al. 2005, 2008; Burgarella et al. 2007; Daddi et al. 2007; Reddy et al. 2006, 2008; Sawicki 2012) and it seems also to be observed at higher redshift (Yabe et al. 2009; Schaerer & de Barros 2010). Other studies (e.g. Bouwens et al. 2009) also suggest the existence of this trend, while the result from Domínguez et al. (2012), revealing also a similar trend up to redshift 1.5 is clearly compatible with our result, but with large uncertain-



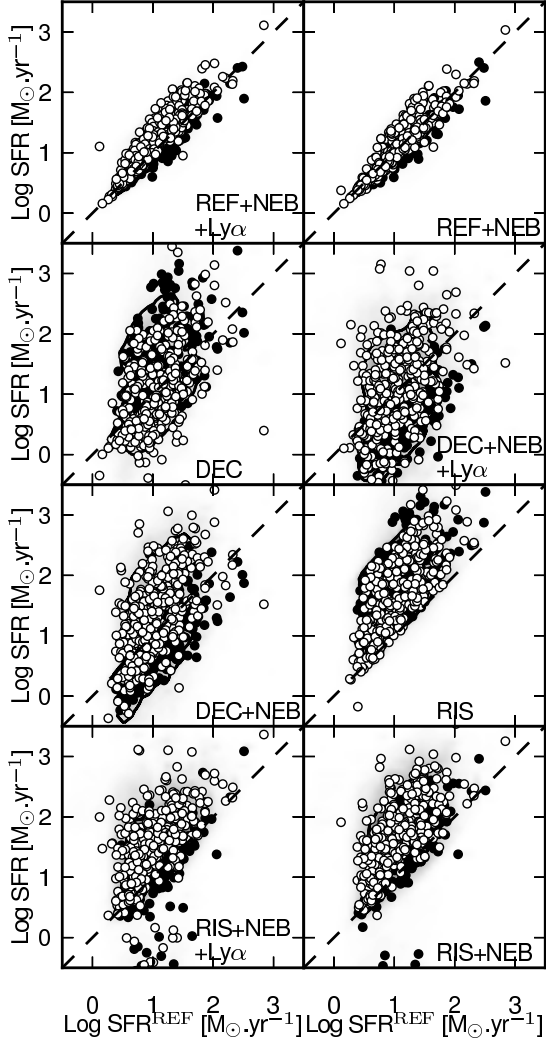
**Fig. 14.** Top: same as Figure 13 for  $A_V$  and  $M_*$ . Bottom: relation between stellar mass and reddening for DEC+NEB model at  $z \sim 4$ . Blue dots for galaxies with median age  $\leq 10^7$  years, red squares:  $10^7 < \text{age} \leq 10^8$ , yellow upward triangles:  $10^8 < \text{age} \leq 10^9$  and black downward triangles: age  $> 10^9$  years

ties in both studies. While this trend could be explained by the age–reddening degeneracy, fixing age at a given value lead to no change. The most natural explanation of this trend is probably that the dust attenuation is related to the stellar mass–metallicity relation (cf. Tremonti et al. 2004; Erb et al. 2006; Finlator et al. 2007; Maiolino et al. 2008).

#### 4.3.7. Star formation rate

The star formation rate (defined here as the instantaneous value at the age  $t$ ) depends quite strongly on the model assumptions, as illustrated in Fig. 15. For constant star formation, the inclusion of nebular emission leads to somewhat higher SFRs on average, due to the younger age, which requires a higher attenuation. The largest differences (up to  $\sim 1$  dex) with respect to constant SFR models are obtained with exponentially declining star formation histories. The reason for such differences is obviously due to the variations in the UV output with time and due to the allowance of young ages ( $< 50$  Myr), which also imply a higher attenuation on average (cf. above).

An interesting feature of the declining SF histories is that it also allows for SFRs lower than those derived using the canoni-

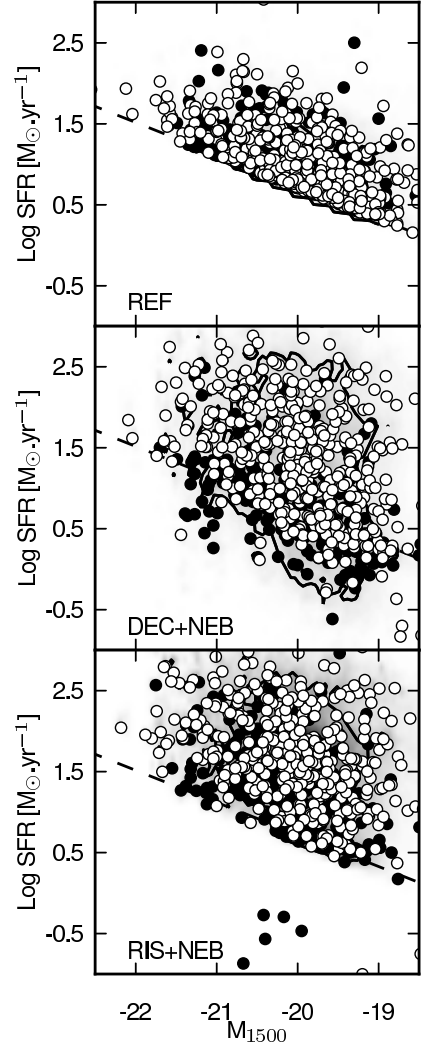


**Fig. 15.** Same as Figure 5 for SFR.

cal calibrations assuming constant SFR. Typically, galaxies with  $t/\tau \lesssim 1$ , have higher SFR (up to 1 dex), if the timescale is short enough to diverge significantly from constant SF. Galaxies with  $t/\tau \gtrsim 2$  are much more quiescent<sup>1</sup> and have lower SFR (up to 1 dex), and at last, intermediate galaxies have SFRs consistent with results from the REF model. This larger “dynamic range” may well be physical, as indicated by the existence of “strong” and “weak” nebular emitters, as we will discuss below. Models with rising star formation histories lead to the highest SFRs, since their SED is always dominated by young stars, which implies a narrower range of UV-to-optical fluxes, hence requiring on average a higher attenuation than for other SF histories (cf. above, Schaerer & Pelló 2005). For rising SF, with ages above  $\sim 10^8$  yr, all galaxies follow the canonical relation (Kennicutt 1998). Below this age, for the same reason as for decreasing SF, the SFR estimated by SED fitting is higher (regardless of dust reddening).

Taking into account nebular emission does not lead to any significant change in the median SFR for constant star formation, while for decreasing SF the median SFRs are lower (mainly

<sup>1</sup> By quiescent galaxies we here mean objects with lower SFR, not galaxies with  $\text{SFR} \sim 0$ .

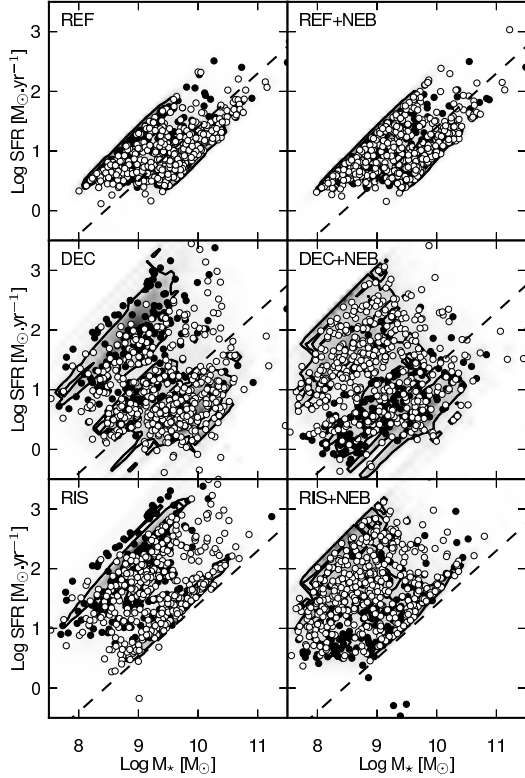


**Fig. 16.** Composite probability distribution of  $M_{1500}$  and SFR for the REF (top), DEC+NEB (centre) and RIS+NEB model (bottom) for the sample at  $z \sim 4$  as determined for each galaxy from our 1000 Monte Carlo simulations. The points overlaid show the median value properties for each object in the sample, black dots for “weak” nebular emitters and white dots for “strong” nebular emitters. The overlaid contour indicate the 68% integrated probabilities on the ensemble properties measured from the centroid of the distribution. The dashed line represents the Kennicutt relation (Kennicutt 1998).

for +NEB+Ly $\alpha$ ) or equal (mainly for +NEB), and for rising SF median SFRs are systematically lower, typically by a factor  $\sim 2$ . From  $z \sim 3$  to 5, this effect is due to difference in dust reddening and age estimations, and in the case of +NEB+Ly $\alpha$  due to the contribution from Ly $\alpha$  line, which can decrease the UV flux necessary to fit the measured fluxes.

Relying on our previous identification of “strong” nebular emitters and “weak” nebular emitters (Section 4.1), we are able to check the consistency of star formation rate estimation. Indeed, since emission lines are produced by the strong UV flux from OB stars in H II regions, we should find a higher SFR for “strong” nebular emitters than for “weak” nebular emitters. As shown in Figure 15, the REF model does not reproduce such a

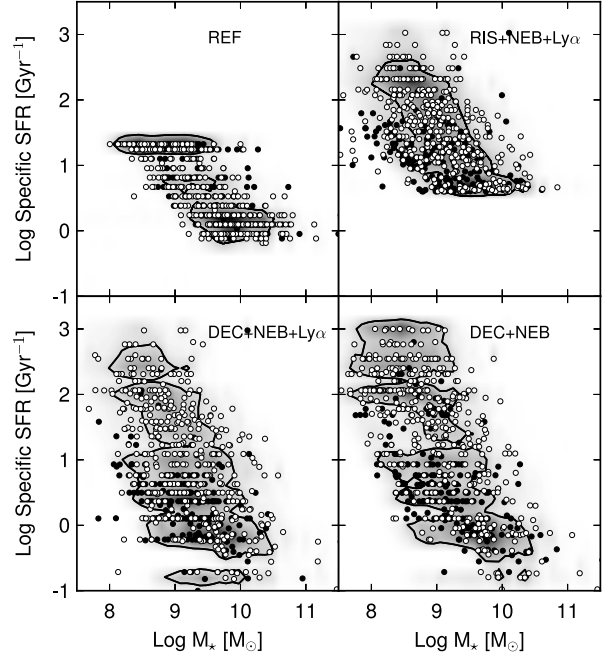




**Fig. 17.** Same as Figure 13 for  $M_*$  and SFR. The dashed line represents the SFR- $M_*$  relation found in Daddi et al. (2007) at  $z \sim 2$ .

separation, since SFRs are roughly similar for both populations or even showing the opposite trend to what is expected. The inclusion of nebular emission with a constant SF does not lead to a significant difference on median SFR estimations between the two populations. Other models without nebular lines (DEC and RIS) do not provide better result than constant SF since they also provide a trend opposite to what is expected: larger median SFR for “weak” nebular emitters. On the other hand, when nebular emission is included the two populations are naturally separated in terms of SFR, as shown in Fig. 15, revealing the “strong” nebular emitters as objects with a strong ongoing star formation episode, and “weak” emitters as a more quiescent population. The capacity of both the declining and rising star formation histories to distinguish these populations can be easily understood, since in this case at least two star formation regime are possible. Separating the two LBG populations we find that the median SFR is higher by  $\sim 0.6$  dex (up to 0.75) for the “strong” nebular emitters from  $z \sim 3$  to  $z \sim 5$  compared to the “weak” emitters, although the typical uncertainty is relatively large ( $\sim 0.5$  dex). At  $z \sim 6$ , only DEC/RIS+NEB+ $\text{Ly}\alpha$  allow to retrieve the expected trend, showing that the presence of just one emission line like  $\text{Ly}\alpha$  can have a large impact on parameter estimation (Schaerer et al. 2011).

We now explore the SFR- $M_{1500}$  relation, illustrated in Figure 16 for the  $z \sim 4$  sample using three different models sets. For the constant star formation (REF) model, the SFR- $M_{1500}$  match with Kennicutt calibration (Kennicutt 1998), taking into account the effect of dust extinction. As explained in Kennicutt (1998), the relation is valid for galaxies with continuous star formation over time scales of  $10^8$  years. SFR/ $L_{1500}$  will be significantly higher



**Fig. 18.** Same as Figure 13 for specific SFR.

in bursty galaxies with a decreasing SF and a short timescale, or simply galaxies younger than  $10^8$  years. This will lead to significantly higher SFR than those given by the Kennicutt relation, regardless of dust reddening. The SFRs found cover a large range of possible values, strongly depending of both SFH and fitting or not with nebular emission (see Table A.1, A.2 and A.3). Again, only rising and declining SF with nebular emission, are able to separate the two populations previously identified as “strong” and “weak” nebular emitters, since they naturally separate these groups into higher and lower SFR galaxies.

The SFR as a function  $M_*$  is plotted in Figure 17 for the  $z \sim 4$  sample. The figure shows that for constant star formation (REF model), we find a relation compatible to that found at  $z \sim 2$  by Daddi et al. (2007) with a relatively small dispersion, and no significant difference if we consider nebular emission. For decreasing SF, our results remain compatible with the relation at  $z \sim 2$  but with a very large dispersion, which can be explained by the large range of timescale, while for rising SF, the star formation rates are systematically higher than those expected from the SFR-mass relation derived at  $z \sim 2$ . We note that for rising and decreasing star formation the galaxies seems to be separated in two groups, where actively star forming galaxies show higher SFRs than expected from the Daddi et al. (2007) relation, and a group of more quiescent galaxies is compatible with this relation. These groups correspond again largely to those previously identified as “weak” nebular emitters and “strong” nebular emitters.

The specific star formation rate (sSFR = SFR/ $M_*$ ) is plotted for  $z \sim 4$  as function of stellar mass in Figure 18. For all models it decreases on average with increasing  $M_*$  and with decreasing redshift. The relation is fairly similar among all models, but decreasing and rising star formation histories lead to higher sSFR values, by more than 1 dex. This increase is significant compared to the typical errors, which range from  $\sim 0.2$  dex for models with constant SF to  $\sim 0.6$  for decreasing and rising SFHs. For decreasing and rising SF, the presence of  $\text{Ly}\alpha$  leads to slightly lower SFR and higher  $M_*$  (except at  $z \sim 6$  where this trend



is reversed), which explains the lower sSFR when compared to models assuming no  $\text{Ly}\alpha$  emission. Comparing declining star formation histories to others yield lower sSFR for some galaxies. Both with the DEC+NEB(+ $\text{Ly}\alpha$ ) and RIS+NEB(+ $\text{Ly}\alpha$ ) models, “strong” nebular emitters have a lower median  $M_\star$  and higher sSFR than the “weak” nebular emitters. This decrease of sSFR with  $M_\star$  can have a physical explanation since the star formation efficiency decreases with stellar mass (eg. Kauffmann et al. 2003).

#### 4.3.8. Metallicity

Metallicity is the parameter least constrained by our SED fits. Indeed, for individual objects the 68% confidence interval for all samples covers basically the three metallicity values (0.02, 0.2, 1  $Z_\odot$ ) used here. Considering the median metallicity there is a trend for RIS+NEB(+ $\text{Ly}\alpha$ ) and DEC+NEB(+ $\text{Ly}\alpha$ ) models to show an increase of the metallicity with galaxy mass. However, the uncertainties are too large to provide firm conclusions. This is consistent with the well known fact that photometry is a poor metallicity indicator.

#### 4.4. Evolution of the physical properties from $z \sim 3$ to $z \sim 6$

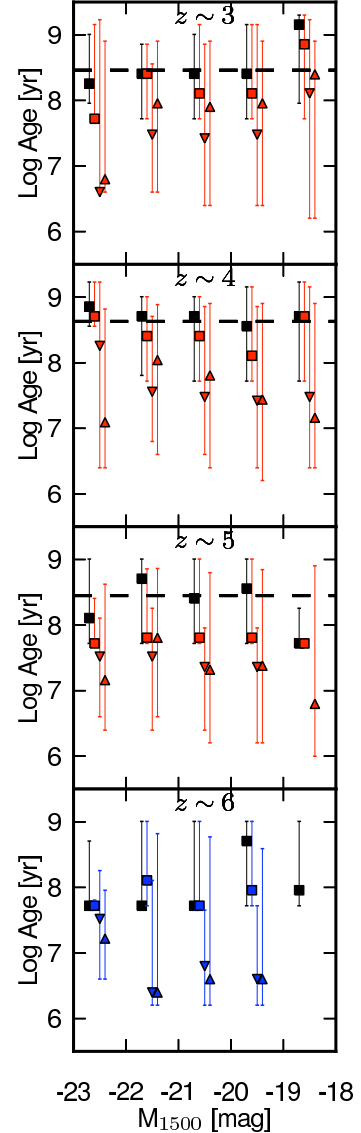
We will now examine the evolution of the median physical properties with redshift and discuss several implications. To allow a meaningful comparison avoiding in particular variations of the completeness limit and of the galaxy luminosity function with  $z$ , we do so in bins of absolute UV magnitude. To discuss the physical parameters derived from models including nebular emission we choose the models with  $\text{Ly}\alpha$  flux set to zero (+NEB option) for  $z \sim 3-5$ , and the NEB+ $\text{Ly}\alpha$  option with maximum  $\text{Ly}\alpha$  emission for  $z \sim 6$ . Within the options discussed in this paper, this choice describes best the trend of increasing strength of  $\text{Ly}\alpha$  with redshift as observed from spectroscopic surveys (eg. Ando et al. 2006; Ouchi et al. 2008; Stark et al. 2010), other studies (Blanc et al. 2011; Hayes et al. 2011), and from our models allowing for varying  $\text{Ly}\alpha$  strength (Schaerer et al. 2011).

In the following, while we plot parameters as a function of  $M_{1500}$ , in bins of magnitude, we notice that at each redshift, there is a very small number of object in the brightest bin, while the faintest bin falls beyond the completeness limit. It is therefore more appropriate to consider mainly the three intermediate bins to examine trends of the physical properties with UV magnitude.

##### 4.4.1. Age

The evolution of median ages with  $z$  is shown in Figure 19. The median age increases with decreasing redshift for the four models, typically by one order of magnitude, with no significant trend with  $M_{\text{UV}}$ . Although declining and rising SFs lead to very similar ages at the highest redshift, they start to diverge somewhat with decreasing redshift. We found that the age differences are essentially driven by the group of quiescent galaxies with “weak” nebular emission. Indeed, fitting active galaxies with rising or decreasing SF leads to very similar ages, while for quiescent galaxies the ages are similar at  $z \sim 6$  to those for rising SF, but start to differ at lower redshift. This is explained both by nebular emission and intrinsically aging population with decreasing redshift. Quiescent galaxies with declining star formation prefer systematically short timescales  $\tau$  (median at 10 Myr at each redshift) to minimize the contribution of nebular emission, which is achieved with ages significantly older than  $\tau$

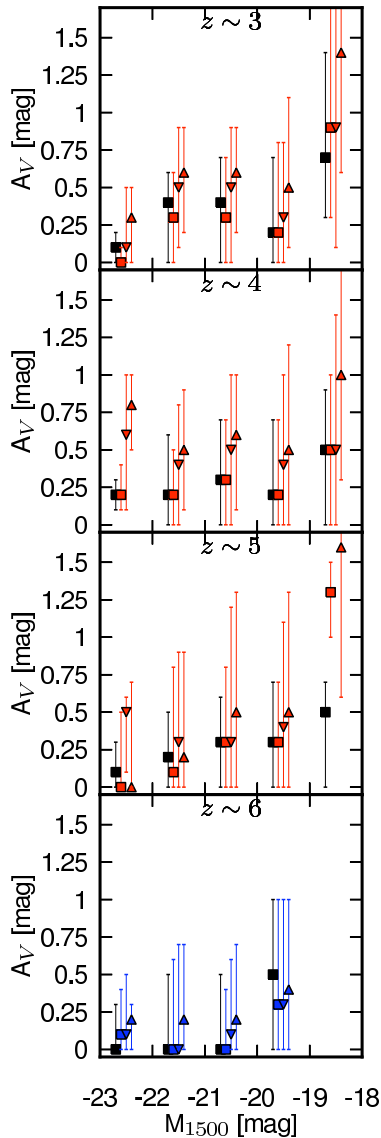
(at least  $t/\tau \gtrsim 2$ ). Indeed, for such short timescales a significant Balmer break can then easily be obtained with only slightly older ages. On the other hand, for rising SF the only way to minimize the contribution of nebular emission (i.e. the equivalent widths of lines) is by choosing relatively old ages. To fit SEDs both devoid of strong emission lines and with a strong Balmer break, even older ages are, however, needed.



**Fig. 19.** Median ages given in bin of absolute magnitude at 1500Å (no correction for dust) from  $z \sim 3$  to 6 for four models. Squares illustrate constant SFR, downward (upward) triangles decreasing (rising) SFHs. Black symbols stand for models without nebular emission (NEB), red symbols for models including nebular emission (NEB), and blue symbols for NEB+ $\text{Ly}\alpha$ . The error bars correspond to the 68% confidence limits of the probability distribution in each bin. Dashed lines show the amount of time spanned since the previous redshift bin.

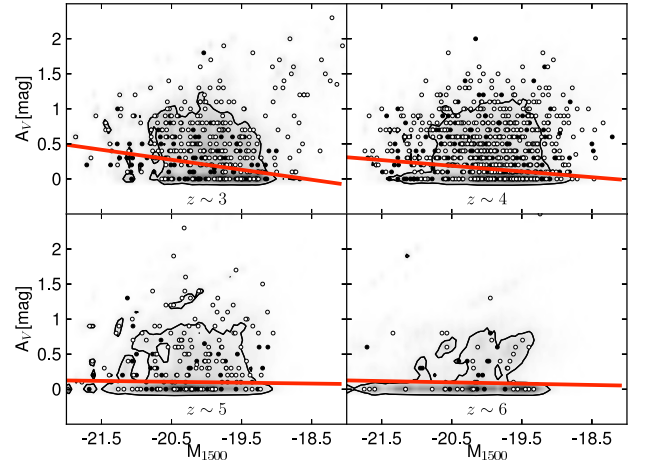
Figure 19 leads one to conclude that for all the models, except the one with constant SFR and no nebular emission (REF), the bulk of the LBG populations at each redshift are dominated

by young galaxies that were not present or visible at the previous redshift. For a constant (REF+NEB) and rising (RIS+NEB) star formation, these results imply that an important number of galaxies formed in the interval from one redshift to another, i.e. in  $\sim 300\text{--}400$  Myr. For decreasing star formation, due to the presence of actively star-forming and quiescent galaxies, we can interpret this result with a scenario with episodes of active star formation which are followed by more quiescent episodes. In this scenario, age becomes a poorly constrained parameter since the underlying older stellar population, formed in previous episodes of strong activity, can be dominated by a newly emerged population. Obviously the *absolute* ages also depend on the SF timescale. For example, fixing the timescale to  $\tau = 100$  (300) Myr for the DEC+NEB models leads to no significant differences on the median values of the stellar mass, SFR or reddening, whereas the median age increases by a factor  $\sim 2$  (3) for both active and quiescent galaxies.



**Fig. 20.** Same as Figure 19 for  $A_V$ .

#### 4.4.2. Reddening



**Fig. 21.** Same as Figure 13 for  $M_{1500}$  and  $A_V$  for DEC+NEB model for  $z \sim 3$  to  $z \sim 5$  and DEC+NEB+Ly $\alpha$  model for  $z \sim 6$ . The red line is a linear fit among the whole probability distribution function.

As shown in Fig. 20, the median dust attenuation decreases with increasing redshift for all the four models considered here. Inclusion of nebular emission for a constant star formation does not provide significant difference, while decreasing and rising star formation histories lead respectively to  $+0.1\text{--}0.2$  and  $+0.3\text{--}0.4$  mag in reddening compared to constant SFR. Some studies (eg. Bouwens et al. 2009; Castellano et al. 2012) shown that there is a trend of decreasing  $\beta$  with increasing absolute UV magnitude, which induce a similar trend on  $M_{1500} - A_V$ , not seen in Figure 20. This trend should be found whatever is the actual  $\beta - A_V$  relation since both classical relation (Bouwens et al. 2009) and the relation found at  $z \sim 4$  in this study (Section 4.3.4) can not change the sign of the slope.

To avoid statistical effects (bins with small number of objects), we use the whole probability distribution at each redshift to determine precisely the relation between  $A_V$  and  $M_{UV}$ , with a result shown in Figure 21. Interestingly, the only way to obtain at each redshift slopes consistent with the relation between  $\beta$  and  $M_{UV}$  (Bouwens et al. 2009; Castellano et al. 2012), is to consider DEC+NEB model from  $z \sim 3$  to  $z \sim 5$ , and DEC+NEB+Ly $\alpha$  at  $z \sim 6$ . For example, rising SF leads to positive slope at  $z \sim 5$ , as constant SF at  $z \sim 6$  (both for any model considered).

As already mentioned in Sect. 4.3.4 the median UV attenuation obtained from our models with rising and declining star formation histories and nebular emission is higher than predicted from conventional methods relying on the average UV slope. For all models, reddening evolves similarly with redshift: it increases with decreasing redshift.

Unfortunately, while photometry in the filters commonly used to determine  $\beta$  (Bouwens et al. 2011) is available for the whole sample at  $z \sim 4$ , this is not the case at  $z \sim 5$  and 6, where less than 20% of the necessary fluxes are available for the  $V$ - and  $i$ -drop samples. To circumvent this shortcoming we use the fluxes predicted by the best-fit model in these filters. At  $z \sim 5$ , using the equations of Bouwens et al. (2011) to derive  $\beta$ , we also find a deviation from the classical  $A_V - \beta$  relation, although less important than at  $z \sim 4$ , with an intermediate relation between the classical relation and that we found, while at  $z \sim 6$ , the result is near from equations 1, 2 and 3 but does not allow us to distin-

guish among them since at low  $A_V$  they are very near and there is only few objects at  $z \sim 6$  with significantly high extinction. We have to notice that we use a  $J$  band from VLT/ISAAC while Bouwens et al. (2011) use data from Hubble/WFC3, which can lead to some differences.

#### 4.4.3. Star formation rate

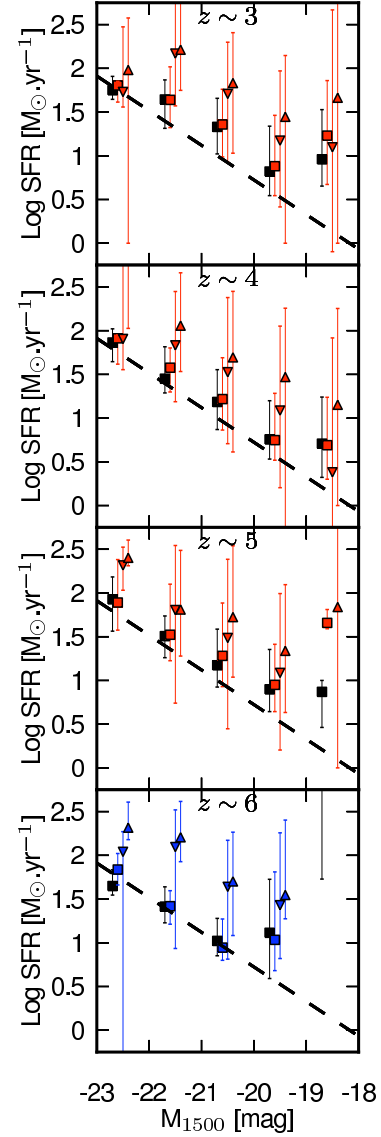
The behavior of the instantaneous SFR as a function of the UV magnitude and for all redshifts is shown in Figs. 22 and 23, where we have separated the sample in the two groups, LBGs with “strong” and “weak” nebular emission (also referred to as ‘active’ and ‘quiescent’ galaxies here), identified earlier. For quiescent galaxies, all the models lead to very similar median SFRs, while for active galaxies, rising and declining star formation histories lead to significantly higher median SFRs, as already discussed. For each group we do not observe a significant change of the SFR with redshifts. The SFR– $M_{UV}$  relation does not evolve with redshift since median age increases with decreasing redshift, while median reddening decreases, so the two effects cancel out on average.

In the next section we discuss some implications from these results on the star formation history.

#### 4.4.4. Stellar mass and implications for star formation history

Stark et al. (2009) suggest to explore the star formation history of LBGs through the evolution of the stellar mass in bins of UV magnitude, since the (observed/uncorrected) UV magnitude is the most reliable tool at high redshift to track star formation, and since stellar mass is the most reliable parameter, i.e. the one least dependent on different assumptions (SFH, metallicity, dust) (Finlator et al. 2007; Yabe et al. 2009). Despite the fact that this study challenges this last assumption (Section 4.3.6), the  $M_\star$ – $M_{1500}$  relation remains a useful tool able to test/falsify some scenarii, mainly constant star formation history. As explained in Stark et al. (2009), if the bulk of galaxies formed stars with a constant star formation rate over sufficiently long time we would expect to see a systematic increase of the normalisation of the  $M_\star$ – $M_{1500}$  relation with cosmic time, i.e. decreasing redshift. If no such change is observed, the SFR cannot be constant, at least not over more than  $\Delta t \gtrsim 300$ – $400$  Myr, the time corresponding to  $\Delta z \approx 1$  between our samples. In other words a non-evolution of the  $M_\star$ – $M_{1500}$  relation with redshift would require other star formation histories or the repeated emergence of new galaxies dominating at each redshift.

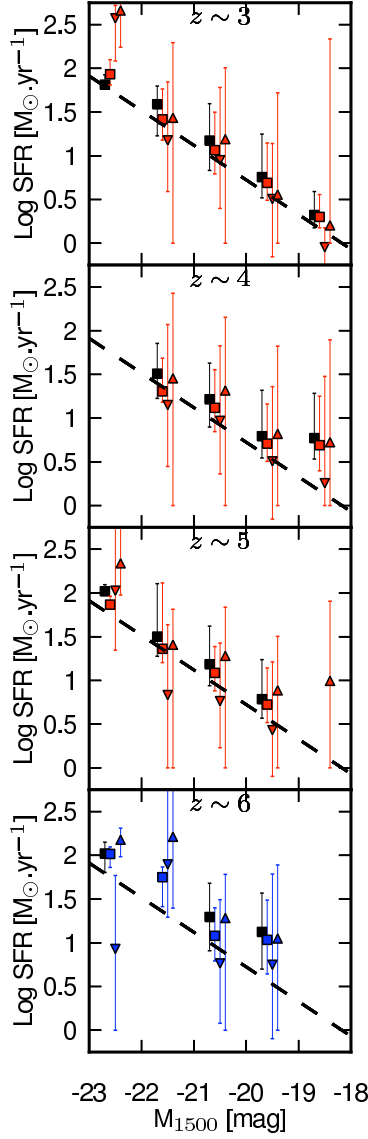
The predicted relation between mass and UV magnitude obtained from all model sets and for all redshifts is shown in Fig. 24. Although obviously the absolute stellar masses depend on the model assumptions, all models yield essentially no evolution of the  $M_\star$ – $M_{1500}$  relation between  $z \sim 5$  to 3, but some shift between  $z \sim 6$  and 5. In particular from redshift 5 to 3, models with constant star formation (both with or without nebular emission, i.e. REF+NEB or REF) do not yield an evolution of stellar mass which would be consistent with the assumption of constant SFR. Even considering the median ages obtained from the fits (Figure 19), it seems difficult to reconcile the picture of constant star formation with the derived parameters. In contrast, the evolution of  $M_\star$ – $M_{1500}$  from  $z \sim 6$  to 5, for REF model, is fully compatible with what is expected from a constant star formation, while for the same model with nebular emission, we still do not see the expected evolution. However, for this latter case, median ages allow to assume that at each redshift, young LBGs domi-



**Fig. 22.** Same as Figure 19 for SFR and for active galaxies. The dashed line shows the Kennicutt relation (Kennicutt 1998).

nate samples. We conclude that constant SF over long timescales is not compatible with the data.

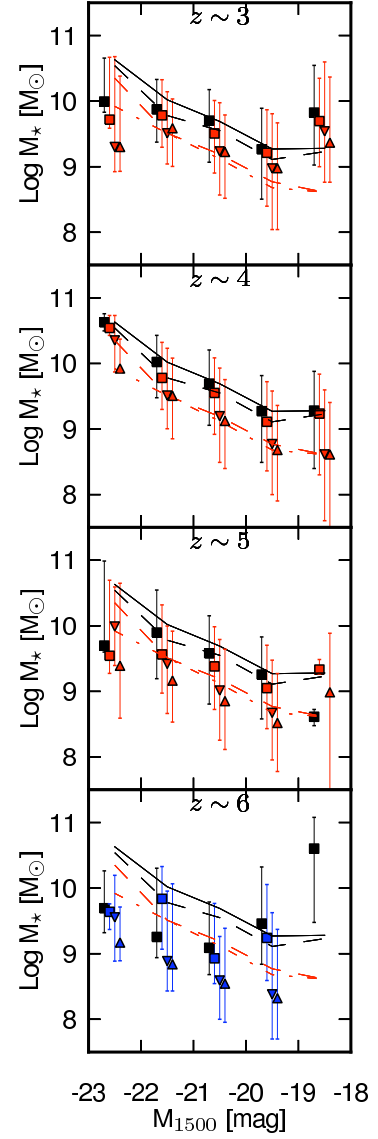
Assuming a constant star formation and no strong evolution of the dust content or its geometric distribution with cosmic time, we expect galaxies to evolve at constant  $M_{UV}$ . With this SFH, the evolution of UV luminosity function should be mainly due to the emergence of new galaxies. For  $M_{UV} = -20.5$ , the number of galaxies per  $\text{Mpc}^3$  increases by a factor  $\sim 3$  between  $z \sim 6$  and  $z \sim 4$  (Bouwens et al. 2007), which means that at least one third of LBGs seen at  $z \sim 4$  must have LBGs at  $z \sim 6$  as progenitors at this magnitude. Indeed, at  $z \sim 4$  and for  $M_{UV} \approx -20.5$ , our sample shows one third of the LBGs having age  $> 700$  Myr, for REF+NEB model, which correspond to the time spans between  $z \sim 6$  to  $z \sim 4$ . We predict the median stellar mass of this galaxies at  $z \sim 4$ , using our parameter estimation at  $z \sim 6$ : the median stellar mass should be  $\approx 10^{10.11} M_\star$ , while our sample of galaxies at  $z \sim 4$  sufficiently older to be descendants of  $z \sim 6$  galaxies have a median stellar mass  $\approx 10^{9.55}$ . This difference of  $\sim 0.6$



**Fig. 23.** Same as Figure 19 for SFR and for quiescent galaxies. The dashed line shows the Kennicutt relation.

dex is three times larger than typical uncertainty on stellar mass estimation with REF+NEB model, leading us to conclude that despite a better consistency of the data with constant SF when nebular emission is taken into account, there is still significant discrepancies.

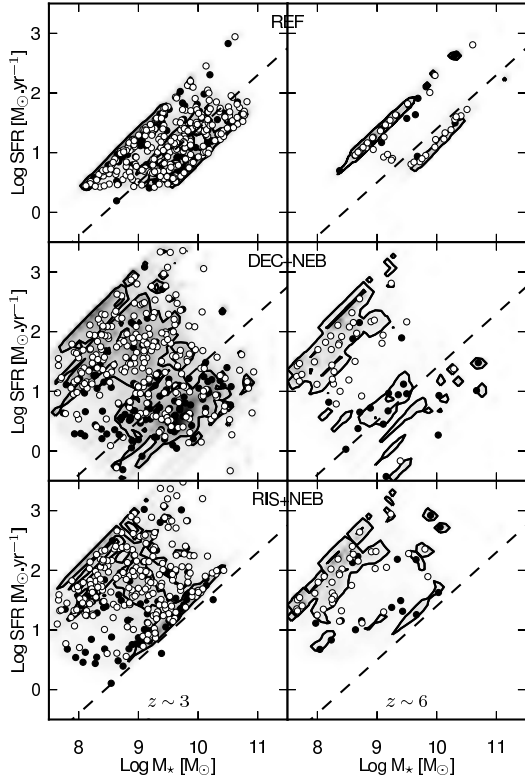
What about the rising star formation history adopted here? It is straightforward to compute how the stellar mass and SFR of each galaxy evolves with time, both forward and backward. Starting from the best-fit for the 1000 MC simulations for each object, we determine how they would evolve in the future. Taking for e.g. LBGs at  $z \sim 5$  which have a median stellar mass of  $10^{8.9} M_{\odot}$  and a median SFR  $\sim 30 M_{\odot} \text{ yr}^{-1}$ , we find that they would have a median stellar mass of  $10^{11.4} M_{\odot}$  and median SFR  $\sim 2000 M_{\odot} \text{ yr}^{-1}$  (corresponding to  $M_{UV} \approx -26$  assuming the Kennicutt relation) after  $\sim 400$  Myr at  $z \sim 4$ . Obviously such extreme objects are not seen, certainly not in large numbers! To hide most of them from the LBG selection a strong increase of dust attenuation with time would be necessary, and in this case it should be fairly easy to find this large populations as IR/sub-mm



**Fig. 24.** Same as Figure 19 for stellar mass. In each panel, we plot the  $M_{*}-M_{1500}$  found at  $z \sim 4$  for more convenient comparison with other studies (Stark et al. 2009; Lee et al. 2011). Black solid line: relationship for REF model, dashed black line: REF+NEB model, dashed red line: DEC+NEB, and dashed-dotted red line: RIS+NEB.

galaxies. More likely this discrepancy implies that the average rising star formation history adopted here from the simulations of Finlator et al. (2011) are not representative for typical  $z \lesssim 6$  galaxies, that the growth does not continue significantly beyond the current ages of the LBGs, or that the growth is predicted to be too fast, or that a combination of these arguments apply. Rising SFHs with varying timescales will be explored in a future paper. In any case, the study of Papovich et al. (2011) shows for the observed number counts at  $z \sim 3-8$  to be compatible with a cosmologically average rising star formation history, the growth of the SFR has to be fairly slow. Their  $\text{SFR}(t) \propto (t/\tau)^{\alpha}$  with  $\alpha = 1.7 \pm 0.2$  and  $\tau = 180 \pm 40$  Myr corresponds to an increase of the SFR by a factor  $\sim 2$  per  $\Delta z = 1$  from redshift 6 to 3, a much slower growth than predicted during the first  $\sim 100-400$  Myr of the star formation history of Finlator et al. (2011).





**Fig. 25.** Same as Figure 17 for three models (Top: REF, centre: DEC+NEB, bottom: RIS+NEB) at  $z \sim 3$  (Left) and  $z \sim 6$  (Right). For  $z \sim 6$ , models with declining and rising SF include  $\text{Ly}\alpha$ .

Hence applying the average rising SFH of Papovich et al. (2011) to our individual galaxies<sup>2</sup> should yield results fairly close to those obtained from our models with constant SFR, including for their inability to reproduce the observed diversity of emission line strengths traced by the  $(3.6-4.5)\mu\text{m}$  color at  $z \sim 3.8-5$ . We therefore conclude that both rapidly and slowly rising star formation histories *over long time scales* ( $\Delta t \gtrsim 100-200\text{Myr}$ ) are not appropriate to describe individual galaxies. Some mechanism turning off star formation or leading to episodic phases appears to be required.

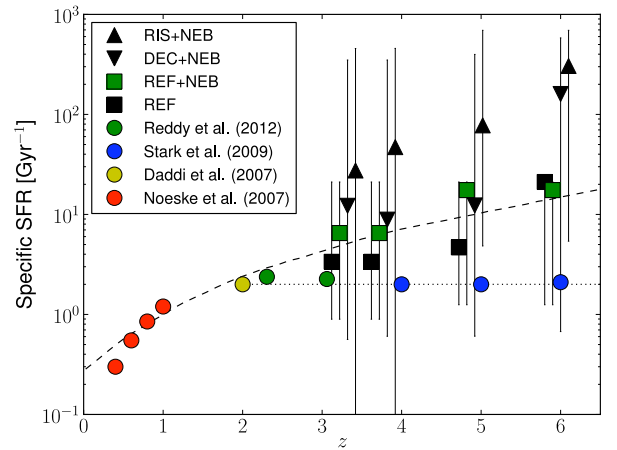
For declining SF, or other episodic star formation histories, it becomes very difficult to connect galaxy populations at different redshifts and to draw conclusions from such a comparison. This is of course due to strong changes in UV luminosity with time, and possibly also to incorrect estimates of some parameters due to the “outshining” effect, where the glare of young massive stars can hide the properties of older stellar population. For example, Papovich et al. (2001) estimate that an hypothetical old stellar population could contain up to  $\approx 3-8$  times the stellar mass of the young stars that dominate the observed SED.

However, Stark et al. (2009) predict, under the assumption of an episodic star formation model, the presence of massive objects with low star-forming activity. Since active and quiescent galaxies have similar distribution in stellar mass from  $z \sim 5$  to  $z \sim 3$ , and significant difference in median SFR ( $\sim 0.6$  dex), we can interpret this result as a confirmation of this latter prediction.

<sup>2</sup> It must be recognized that the average SFH derived by Papovich et al. (2011) corresponds to a cosmologically averaged history, which a priori does not apply to individual galaxies.

A consequence of episodic SF is that age becomes a poorly constrained parameter. It is also difficult to place constraints on the timescale of activity and inactivity, and indeed, to determine if estimated parameters at different redshift are consistent with this scenario. Lee et al. (2009a) provide a higher limit on star formation activity of 350 Myr, based on observed UV luminosity function and clustering at  $z \sim 4$  and  $z \sim 5$ , and Wyithe & Loeb (2011) found that at high-redshift ( $z \gtrsim 6$ ) starburst timescale is set by the lifetime of massive stars, by comparing different assumptions on supernova feedback and observed evolution in galaxy size and UV luminosity function. These two results promote short timescales, which seems to be confirmed by our results, while uncertainties remain large. Furthermore, the result of Wyithe & Loeb (2011) implies decreasing timescale with increasing redshift. Figure 25 show the difference in SFR- $M_*$  relation between  $z \sim 3$  and  $z \sim 6$ : at  $z \sim 6$ , active and quiescent galaxies form two clearly separate groups while at  $z \sim 3$ , the two groups populate all the intermediate state. We can interpret this result with an evolving star formation timescale, with at  $z \sim 6$  a shorter timescale in comparison with  $z \sim 3$  leading to galaxies evolving very rapidly from an active to a quiescent state.

Age becomes more or less a tracer of recent star formation activity and do not allow to determine if observed high redshift LBGs are progenitors of observed low redshift LBGs. In this scenario, the number of free parameter remains too large to conclude: duration of star-formation activity, duration of inactivity, the possibility for LBGs to evolve into a state with no star formation activity. Furthermore, some studies (Bouché et al. 2010; Wuyts et al. 2011) suggest other SFHs, like exponentially increasing or delayed star formation. These scenarios will be studied in Schaerer et al. (2012).



**Fig. 26.** Median specific star formation rate as a function of redshift for four models, with 68% confidence limit based on the whole probability distribution function, with comparison with results from different studies at  $M_* = 10^{9.5} M_{\odot}$  (Noeske et al. 2007; Daddi et al. 2007; Stark et al. 2009; Reddy et al. 2012). At  $z \sim 6$ , we show results with +NEB+ $\text{Ly}\alpha$  option for declining and rising SFHs. Typical errors are  $\sim 0.3$  dex. The dashed line shows the relation expected from Bouché et al. (2010) for an exponentially increasing star formation at a fixed  $M_* = 10^{9.5} M_{\odot}$ . The dotted line, given by  $\text{sSFR} = 2$  is shown to guide the eye.



#### 4.5. Specific star formation rate

Since our study gives some elements supporting an episodic star formation as scenario for star formation history at high redshift, we compare the evolution of sSFR with results from other studies, using the compilation from González et al. (2010), given at fixed stellar mass  $M_\star = 5.10^9 M_\odot$  (Noeske et al. 2007; Daddi et al. 2007; Stark et al. 2009). Since our different models lead to significative different median stellar masses (for e.g.,  $10^{9.6} M_\odot$  for REF model at  $z \sim 3$  and  $10^{8.6} M_\odot$  for DEC+NEB+Ly $\alpha$  model at  $z \sim 6$ ), we use for comparison median sSFR for whole sample at each redshift. Typically other studies found no significative change in median stellar mass with redshift and the median value of sSFR at  $M_\star = 5.10^9 M_\odot$  is near from the median value for the whole sample (e.g. González et al. 2010). Due to completeness limit, these values can be considered as lower limit since we observe a trend of increasing sSFR with decreasing stellar mass for all models, albeit this trend is moderate with constant star formation (see Figure 18). However, we can not exclude the presence at these redshift of high star forming galaxies enshrouded with dust, so our results have to be considered with appropriate caution. Our results are shown in Figure 26.

While our results with a constant star formation seems to be consistent with Stark et al. (2009) results for  $z \sim 4$  and  $z \sim 5$ , we find a higher median sSFR at  $z \sim 6$ . Looking at IRAC channels, galaxies at  $z \sim 6$  with high sSFR ( $\sim 20 \text{ Gyr}^{-1}$ ) have on average 2 to 3 channels with no detection (or no data), while galaxies with lower sSFR (median  $\sim 1.7 \text{ Gyr}^{-1}$ ) have typically no more than 1 channel with no detection. No detection of LBGs in optical (rest-frame) leads to lower stellar mass estimation ( $\sim 1$  dex), while estimated SFR stays similar. This explains higher sSFR found at  $z \sim 6$  with constant star formation. The inclusion of nebular emission to constant star formation (REF+NEB) leads to higher sSFRs, and evolution compatible with an exponentially increasing star formation proposed by Bouché et al. (2010).

Our results with declining and rising star formation including nebular emission differ significantly from previous studies, with a higher median sSFR and a possible trend of increasing sSFR with redshift. While studies without considering nebular emission lead to conclude that star formation seems to be drive by different principles below and above  $z \sim 2$ , our results seem that star formation history can be more homogeneous among cosmic time than suggested by previous studies and models which not include nebular emission.

## 5. Discussion

### 5.1. Do we obtain realistic ages?

The dynamical timescale is a lower limit for age estimation of a galaxy for reasons of causality. Dynamical timescale is  $t_{\text{dyn}} \simeq 2r_{\text{hl}}/\sigma$  and using different studies (Bouwens et al. 2004; Ferguson et al. 2004; Douglas et al. 2010) providing  $r_{\text{hl}}$  values and Pettini et al. (1998) providing velocity dispersion at  $z \sim 3$ , and assuming no evolution of  $\sigma$ , we estimate that  $t_{\text{dyn}}$  evolves from  $\sim 40 \text{ Myr}$  at  $z \sim 3$  to  $\sim 20 \text{ Myr}$  at  $z \sim 6$ . Dynamical timescale is also found to scale with  $(1+z)^{-3/2}$  (e.g. Wyithe & Loeb 2011), which implies a variation of a factor  $\sim 3$  between redshift 3 to 6.

Since we have not imposed a lower limit for age estimation for both declining and rising SF, both models lead to a significant number of galaxies with age below typical dynamical timescale, mainly for active galaxies. Indeed median ages for this latter population is almost always close to the dynamical

limitation in age. Since age estimation is dependent from star formation timescale for declining SF, we test some fixed values of  $\tau$  to examine effect on age and other parameters estimation. With  $\tau = 100 \text{ Myr}$ , we do not observe any significant change in median values of the different parameters, except for age, which increases by a factor  $\sim 2.5$  for both active and quiescent galaxies, at each redshift. However, looking at the age distribution, there is still a significant number of galaxies with age  $< t_{\text{dyn}}$ . Imposing a lower limit  $\sim 40 \text{ Myr}$  leads to an explanation of the previous result: nebular emission of active galaxies seems to be correctly fitted only with very young ages. Indeed, using our sample at  $z \in [3.8, 5]$ , with  $3.6\mu\text{m}$  and  $4.5\mu\text{m}$  fluxes measured, both declining and rising models are not any more able to produce significantly better fit of  $3.6\mu\text{m}$ - $4.5\mu\text{m}$  color (in comparison with REF model), which is correlated to EW(H $\alpha$ ). Associated mean  $\chi_r^2$  values (defined in the label of Figure 4) reach 1.4 and 1.2, respectively for DEC+NEB and RIS+NEB models, while the quality of the entire SEDs remain almost unchanged.

While this discrepancy between a significant fraction of our estimated ages and dynamical timescale can lead to question our study, we remind two elements. First, declining and rising models lead to higher uncertainty on age, typically of a factor  $\sim 3$  when nebular emission is taken into account, which can allow to reconcile almost 80% of our samples with dynamical timescale. Second, there is no measured velocity dispersion of nebular lines at  $z > 3$ , since this measurement is confronted to the limit of current facilities. Furthermore, we use typical velocity dispersion measured at  $z \sim 3$  ( $\sim 80 \text{ km.s}^{-1}$ ) from Pettini et al. (1998), but one object of this study (among five) shows a velocity dispersion more than two times larger than this latter value, which also decreases  $t_{\text{dyn}}$  by a factor two for this object.

We conclude that apparent discrepancies between dynamical timescale and estimated ages in this study are considerably reduced when uncertainties on age estimation, and observations limitations are taken into account.

### 5.2. Comparison with other studies

During the last decade, various papers provide analysis of the properties of LBGs properties at high redshift. Since we use a large range of star formation histories, we are able to compare our results with several of them, carried out for various LBG samples between  $z \sim 3$  and 6. The main assumptions made for the SED fits in these papers, the size of their galaxy samples, and our corresponding model for comparison are listed in Table 2. To allow straightforward comparisons we will use the median/mean values derived from our probability distribution functions. Except stated otherwise, we determine the mean over our entire samples. Although derived from the same pdfs, the values from our models quoted here do therefore not correspond to values listed in Tables A.1–A.3 (which are *median* values).

Shapley et al. (2001, hereafter S01) and Papovich et al. (2001, hereafter P01) have studied 74 and 33 LBGs respectively, all at  $z \sim 3$ . The sample of S01 is restricted to brighter objects ( $M_{\text{UV}} \lesssim -20.7$ ) than our  $z \sim 3$  sample, so we compare their results only with the results for a subsample of our objects with the same magnitude limit. Note that both studies include photometry up to the K-band, but not beyond. For both studies, we find very similar results for the median ages (S01 obtain  $\sim 350$ , P01  $\sim 70 \text{ Myr}$ ) and the age distribution when the corresponding models without nebular emission are used. This is not surprising since age is mainly constrained by the presence of the Balmer Break, which both studies can constrain from K<sub>S</sub> and J (or H) band data. We also find a similar median stellar mass

**Table 2.** Summary of other studies in the literature. Col. 3 indicates their assumed star formation histories, col. 4 the extinction law, col. 5 the number of galaxies, and col. 6 indicates our model used for comparison.

Redshift	Authors <sup>a</sup>	SFH <sup>b</sup>	ext. law	$N$	Comparison model
$z \sim 3$	S01	constant SFR	Calzetti	74	REF
$z \sim 3$	P01	exp. declining	Calzetti	33	DEC
$z \sim 4$	PG07	exp. declining	Calzetti & SMC	47	DEC
$z \sim 4$	L11	constant SFR	Calzetti	6 <sup>c</sup>	REF
$z \sim 5$	S07	SSP/constant/exp. declining	Calzetti	14	DEC
$z \sim 5$	V07	constant SFR	SMC	21	REF
$z \sim 5$	Y09	constant SFR	Calzetti	105	REF
$z \sim 6$	Y06	SSP/constant SFR	Calzetti <sup>d</sup>	53	DEC
$z \sim 6$	E07	SSP/constant/exp. declining	Calzetti	17	DEC

<sup>a</sup> S01: Shapley et al. (2001), P01: Papovich et al. (2001), PG07: Pentericci et al. (2007), L11: Lee et al. (2011), S07: Stark et al. (2007), V07: Verma et al. (2007), Y09: Yabe et al. (2009), Y06: Yan et al. (2006), and E07: Eyles et al. (2007).

<sup>b</sup> SSP: single stellar population, burst.

<sup>c</sup>: stack of 1913 objects in 6 UV magnitude bins.

<sup>d</sup> Yan et al. (2006) assume  $A_V = 0$ . For a fair comparison, we run DEC model with the same assumption.

compared to P01 ( $3.10^9 M_\odot$ ), while we obtain mean stellar mass lower by  $\sim 0.2$  dex compared to S01 ( $2.10^{10} M_\odot$ ). Since P01 provide best-fit parameters for their sample, we can see that significant discrepancies appear on dust attenuation, with for both studies extinction factors  $\sim 2$  times larger in comparison with our results. A plausible explanation is that since longer wavelengths are less sensitive to reddening, we can better constrain this parameter with IRAC data. Furthermore, S01 used a BC96 population synthesis which can lead to increase reddening up to 0.2 mag in  $A_V$  in comparizon with CB01. However, S01 also indicated that this difference in population synthesis should lead to significantly older ages, what we do not see.

Pentericci et al. (2007, hereafter PG07) and Lee et al. (2011, hereafter L11) provide analysis of 47 and 1913 LBGs at  $z \sim 4$ , with individual SED fitting in the first case, while L11 used a stacking procedure in UV magnitude bins. The large sample of L11 has a lower  $M_{UV} = -21.43$ , so as at  $z \sim 3$ , we take care to do an appropriate comparison. Our results are similar with those from L11 in stellar mass, age and dust extinction, while we have only 17 LBGs with  $M_{UV} \leq -21.43$ . A discrepancy appears with the results from PG07, since we find a lower mean stellar mass by  $\sim 0.4$  dex, while age, dust extinction and SFRs are similar.

Stellar masses at  $z \sim 5$  and 6 are also found in good agreement with previous studies, when the same/similar assumptions are used. Concretely, at  $z \sim 5$ , V07 found a median stellar mass of  $2.10^9 M_\odot$  ( $3.6.10^9 M_\odot$  for our study), and Y09 and S07 found a mean stellar mass of  $4.1.10^9$  and  $7.9.10^9 M_\odot$  respectively (while we find  $9.9.10^9$  and  $1.2.10^{10} M_\odot$  with REF and DEC model respectively). At  $z \sim 6$ , Y06 and E07 found  $9.6.10^9$  and  $1.6.10^{10} M_\odot$  respectively, compared to our masses of  $9.9.10^9$  and  $1.1.10^{10} M_\odot$ . The consistency of our results with other studies, taking into account typical uncertainty of  $\sim 0.15$  dex for REF model and  $\sim 0.2$  dex for DEC model, confirm that stellar mass is the most reliable parameter, since different assumptions on star formation history and also extinction law lead to similar results (cf. Papovich et al. 2001; Verma et al. 2007; Yabe et al. 2009).

Ages of  $z \sim 5$ –6 LBGs obtained in the literature agree overall with our results, except for the study of Verma et al. (2007). Indeed, compared to their young median age of 25 Myr we find 255 Myr, a difference which cannot be explained by uncertainty ( $\sim 0.15$  dex for REF model and  $\sim 0.4$  dex for DEC model). A possible explanation for the discrepancy with V07 could be the age-reddening degeneracy. However, V07 found a median  $A_V = 0.3$  mag, while we find 0.2 mag, which cannot explain the

difference on the median age estimate. The origin of this difference remains unclear. On the other hand, Y09 and S07 found median ages of 25 Myr and 288 Myr (compared to 52 Myr and 320 Myr from our study) at  $z \sim 5$ , and Y06 and E07 obtain 290 Myr and 400 Myr (262 Myr and 190 Myr) at  $z \sim 6$ , in good agreement with our results when the same assumptions are made.

Other difference appears on the reddening estimates: while Y09 found a mean  $A_V \sim 0.9$  mag at  $z \sim 5$ , we obtain 0.4 mag for the corresponding constant SFR model (typical uncertainty  $\sim 0.1$  mag). At  $z \sim 6$  E07 found no reddening on average, which we obtain  $A_V = 0.4$  for the DEC model (typical uncertainty  $\sim 0.25$  mag). Our result at  $z \sim 5$  with the DEC model (mean  $A_V = 0.4$ ) is consistent with S07, who found the same mean reddening. The differences on dust reddening estimation can have several explanations: the limited size of the samples in comparison with ours, or the lack of deep NIR and IRAC photometry able to put stronger constraints both on reddening and age. However, these latter arguments can not be used to explain the discrepancy with Y09, and we did not found a satisfactory explanation.

These discrepancies on reddening must obviously lead to differences on the derived star formation rates. Indeed, at  $z \sim 5$  Y09 found a mean SFR of  $141 M_\odot \text{yr}^{-1}$  and V07 found a median SFR of  $40 M_\odot \text{yr}^{-1}$ , while we find  $\sim 50 M_\odot \text{yr}^{-1}$  and  $15 M_\odot \text{yr}^{-1}$  from the models corresponding best to their assumptions. The difference on reddening estimation with Y09 explains the difference on SFR estimation, while the result of V07 can be explained by their very young median age, which leads to a significant deviation from the Kennicutt relation (Kennicutt 1998), and thus to a higher SFR. S07 do not provide SFR estimations but Y09 fits the parameters of S07: they found a median value of  $20 M_\odot \text{yr}^{-1}$ , consistent with our result. However, these refitted parameters differ significantly from original the result of S07 for both stellar mass and age, which casts some doubt on the homogeneity of this comparison. For  $z \sim 6$ , Y06 and E07 found mean SFR  $\leq 10 M_\odot \text{yr}^{-1}$ . While our modeling with similar assumptions as Y06 lead to a similar result ( $7 M_\odot \text{yr}^{-1}$ ), our SFR estimation differs significantly from E07, with an higher SFR ( $\sim 80 M_\odot \text{yr}^{-1}$ ). This latter difference can be explained by the higher estimated dust reddening, since E07 estimated that  $z \sim 6$  galaxies are mainly dust free.

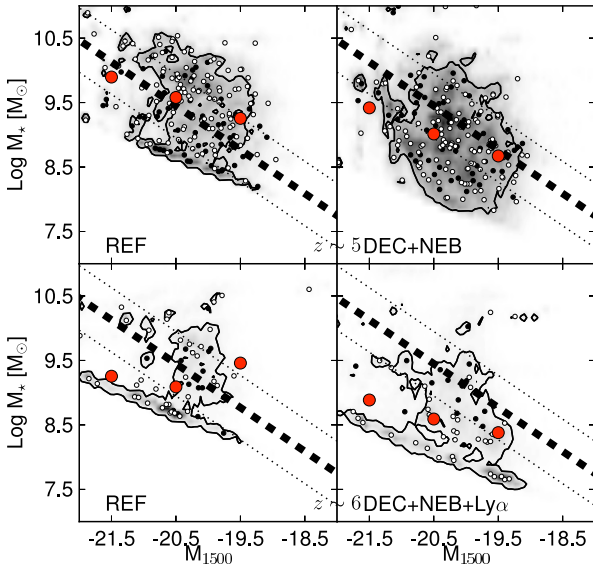
Overall, for identical assumptions, we find good agreement with other studies in general (exception: young ages of V07 and high dust extinction of Y09). But when nebular emission is in-

cluded, the results change quite significantly, as we have shown here.

### 5.2.1. Evolution of the mass – UV magnitude relation?

Stark et al. (2009) and González et al. (2011) cover a range of redshift between  $z \sim 4$  to  $z \sim 6$  (up to  $z \sim 7$  for the latter). Both studies present the  $M_\star$ – $M_{UV}$  relation, which can directly be compared with our results in Figure 12 and 24. As already discussed in Sect. 4.3.6 we find quite similar results as González et al. (2011), although our model for exponentially declining and for the rising star formation history suggest somewhat lower masses at high UV luminosities. Overall our relations remain, however, within the scatter of  $\pm 0.5$  indicated by their work.

Both Stark et al. (2009) and González et al. (2011) find no evolution of the mass – UV magnitude relation with redshift. In contrast, our results seem to indicate a change of the  $M_\star$ – $M_{UV}$  relation between  $z \sim 5$  and 6, as shown in Fig. 27. (cf. also Sect. 4.4.4). The main reason for this change is due to our finding of relatively young ages for  $z \sim 6$  galaxies, which implies a lower  $M/L_{UV}$  ratio. This effect was also noticed by McLure et al. (2011) for their  $z \sim 6$ –8 sample. A preliminary reanalysis of their sample with the same method and assumptions used in the present paper appears to confirm our finding from the  $z \sim 6$  sample.



**Fig. 27.** Composite probability distribution of  $M_{1500}$  and  $M_\star$  for REF and DEC+NEB(+Ly $\alpha$ ) models at  $z \sim 5$  and  $z \sim 6$ . The black dashed line represents the  $M_\star$ – $M_{1500}$  trend found by González et al. (2011) and black dotted lines shows a scatter of  $\pm 0.5$  dex. The points overlaid show the median value properties for each object in the sample, black dots for “weak” nebular emitters and white dots for “strong” nebular emitters. The overlaid contour indicates the 68% integrated probabilities on the ensemble properties measured from the centroid of the distribution. Red dots are median values of  $M_\star$  in  $M_{UV}$  bins.

### 5.2.2. LBGs with strong emission lines

An important finding from our quantitative modeling of LBGs with spectral models including nebular lines is the distinction of two separate categories of galaxies identified as “strong” and “weak” nebular emitters (cf. Sect. 4.1). Our work has revealed these categories from comparisons of the fit quality for models with and without nebular emission, and interestingly we found approximately the same fraction of objects ( $\sim 2/3$  strong versus  $1/3$  weak emitters) at each redshift.

As already mentioned above, two other studies of  $z \sim 4$ –5 LBGs have previously found similar objects with strong H $\alpha$  emission identified by their excess in the  $3.6 \mu\text{m}$  filter with respect to  $4.5 \mu\text{m}$ . Indeed, out of a sample of 74 LBGs with spectroscopic redshifts between 3.8 and 5 Shim et al. (2011) found at least 65% of galaxies showing a  $3.6 \mu\text{m}$  excess attributed to H $\alpha$ . Earlier Yabe et al. (2009) had already noted a  $3.6 \mu\text{m}$  excess for 70% of their  $z \sim 5$  LBG sample of  $\sim 100$  galaxies, attributed to the same effect. Obviously our study finds the same result and a very similar percentage of galaxies. Our work carries this result further by revealing the existence of these two LBG categories among all the samples studied here, ranging from U-drops to i-drops (i.e.  $z \sim 3$ –6), and by extending this result to fainter objects, for which no spectroscopy is currently available.

Besides this import agreement, our results differ, however, on several points from those of Shim et al. (2011). For example, these authors conclude from a comparison of the estimated H $\alpha$  equivalent width (obtained from the  $3.6 \mu\text{m}$  excess) and ages from broad-band SED fits, that 60% of their so-called H $\alpha$  emitters are forming stars at a relatively constant rate, whereas the rest prefer more “bursty” star formation. Our models yield nearly opposite results (cf. Sect. 4.2). Since the effects of nebular emission affect in particular the estimated ages, the models of Shim et al. (2011), which neglect these effects in their SED modeling, cannot give consistent physical parameters for their strong H $\alpha$  emitters. Similarly, whereas Shim et al. (2011) conclude for a preference for the SMC extinction law, our calculations for the B-drop sample using this law show that the vast majority of objects is better fit with the Calzetti attenuation law, when nebular lines are taken into account. Finally, from our models we see no need for a top-heavy IMF or extremely low metallicities, suggested as possible causes of the strong H $\alpha$  emission by Shim et al. (2011). In any case, the H $\alpha$  star formation rates and the H $\alpha$  equivalent widths derived Shim et al. (2011) are comparable to ours.

## 5.3. Remaining uncertainties

### 5.3.1. Uncertainties affecting individual objects

Our study illustrates in detail the way various physical parameters depend on model assumptions made for the broad band SED fits, and a large range of parameter space has been explored here. Despite these extensive investigations, the impact of some assumptions have not been explored in depth, and several uncertainties remain.

The impact of different extinction laws, for example, has hardly been discussed here. Several papers presenting SED fits of LBGs at different redshifts have examined e.g. the differences obtained with the Calzetti attenuation law (adopted here) and the SMC extinction law of (Prevot et al. 1984; Bouchet et al. 1985). Among them are e.g. the work of Papovich et al. (2001); Verma et al. (2007); Yabe et al. (2009) who study LBG samples at  $z \sim 3$  and 5. A comprehensive comparison of the impact of different

extinction laws (including SMC, LMC, Galactic, and the Calzetti law) is presented in Yabe et al. (2009). To examine how different extinction laws modify the results from SED fitting models including nebular emission we have modeled the B-drop sample with the declining star formation history (DEC+NEB model) adopting the SMC law of (Prevot et al. 1984). Qualitatively our results show the same trends as found by Yabe et al. (2009), i.e. no big difference on stellar masses, lower  $A_V$ , older ages, and lower SFR. However, we find that the Calzetti attenuation law provides a better fit for the vast majority of galaxies. Except for few special cases such as the lensed galaxies cB58 and the Cosmic Eye studied in detail and at different wavelengths including the IR (Baker et al. 2004; Siana et al. 2008), where the SMC law appears to be favored e.g. from the measurement of the IR/UV luminosity and the UV slope, it is generally thought that the Calzetti attenuation law is applicable (at least on average) to high redshift star forming galaxies.

Our models including the effects of nebular emission assume case B recombination to compute the strength of the hydrogen recombination lines (and empirical line ratios for other lines from He and metals) and of nebular continuum emission. In particular we do thereby assume that all Lyman continuum photons contribute to nebular emission, neglecting therefore possible losses of ionizing photons due to dust inside HII regions or to escape of Lyman continuum photons. Furthermore, we assume the same extinction/attenuation for stellar and nebular emission, whereas observations of nearby galaxies indicate that emission lines usually suffer from a higher extinction than the stellar continuum (cf. Calzetti et al. 2000). In this sense our models maximize the effects of nebular emission, whereas in reality the lines could be weaker than predicted by our models. On the other hand we have adopted empirical line ratios compiled by Anders & Fritze-v. Alvensleben (2003) from nearby galaxies, whereas the conditions in distant galaxies could be different, leading e.g. to higher excitation or stronger lines (Erb et al. 2010). These uncertainties are currently difficult to quantify. In any case, the broad band SEDs reveal quite clearly the presence of nebular lines and our models bracket a wide range between maximum and no nebular emission. Future spectroscopic observations or narrow/intermediate band imaging should try to provide more detailed and accurate observational constraints on the emission line strengths of LBGs at high redshift. Detailed predictions from our models regarding individual lines will be presented elsewhere and can be made available on request.

As clearly shown by our study, assumptions on the star formation history have a significant impact on the estimated physical parameters when broad band SED fits are used. Although certain star formation histories are found to provide better fits than others (cf. Sect. 4.2), it is obvious that the simple parametrisations commonly adopted in the literature and in our study can only be very crude representations of the true SF histories of individual galaxies. In a companion paper (Schaerer et al. 2012) we have explored two additional families of star formation histories, so-called delayed star formation ( $\text{SFR} \propto t \exp(-t/\tau)$ ), and exponentially rising histories with variable timescales. Using arbitrary star formation histories is, however, not practical, also given the limited number of observational constraints. Another approach has e.g. been taken by Finlator et al. (2007), who have used the star formation history from their cosmological hydrodynamic simulations to fit observed high- $z$  galaxies. However, such studies have so far been limited to a very small number of galaxies. Future improvements on this issue, both from observations and simulations, are certainly needed. E.g. Reddy et al. (2012) test different SF histories using IR and UV observations

of  $z \sim 2$  LBGs. Along similar lines we show in Schaefer et al. (2012) how SFHs can be distinguished by measuring their dust emission with future ALMA observations and/or with measurements of emission lines.

### 5.3.2. Biases and selection effects

For obvious reasons selection effects and various biases affect in general studies of galaxy populations, and these need to be taken into account in comparisons between different samples, in the analysis of apparent correlations between derived physical parameters, and in other contexts.

To allow meaningful comparisons with other studies we have presented our detailed results in bins of UV magnitude (see data in Tables A.1, A.2, and A.3). Especially in the brightest and faintest bins the number of galaxies is quite low, so that these results should be taken with care. Of course, the number of galaxies at different redshifts varies quite strongly, affecting also the accuracy of the median physical properties (and confidence range) derived here.

As in most literature studies, selection effects and biases have not been treated here. The impact of biases on the determination of physical parameters of LBGs from broad-band SED fitting has e.g. been studied by Lee et al. (2009b), who construct mock galaxy catalogs from semi-analytical models, which are then used to fit the simulated galaxies with a standard SED fitting tool. They find that stellar masses can be recovered quite well, whereas single-component SED fitting methods underestimate SFRs and overestimate ages. The differences are attributed in part to a “mismatch” of star formation histories between their fitting tool (which assumes exponentially declining SF) and those predicted by their galaxy models (which are often rising). A similar “template mismatch” was also identified as the main cause for differences in the comparison of  $z \sim 1.5$ –3 merger simulations with SED fitting results carried out by Wuyts et al. (2009). Our models are also prone to such biases/uncertainties, and the role of the assumed SF histories on the derived physical parameters has already been discussed above. However, since we have covered a wider range of SFHs including rising SF, our results may less suffer from this potential problem. The results from SED fits with additional SFHs (including exponentially declining histories with adjustable timescales) are presented in Schaefer et al. (2012).

Various correlations have been found between physical parameters as the stellar mass, the star formation rate, age and others, both in the literature and in our study. Although our work emphasizes in particular the way the physical parameters depend on various model assumptions, i.e. a relative approach, it is important to be aware of the selection effects which may significantly affect such correlations. Stringer et al. (2011), e.g. have recently examined the behavior of the SFR and the specific SFR with stellar mass, two important quantities discussed extensively in the recent literature. From their simulations of mock galaxies to which observational selection criteria and “standard” analysis are applied, they show how true underlying trends can be misrepresented. This study also echoes the caution expressed by Dunne et al. (2009) on the apparent sSFR–mass relation, which they urge, could be severely affected by selection biases. If and to which extent apparent correlations between extinction and mass, and age and mass (cf. Sect. 4.3.6) may be affected by selection effects and biases has not yet been addressed, to the best of our knowledge. A quantitative study of these effects is clearly beyond the scope of the present publication.

## 6. Summary and conclusions

We present an homogeneous study of a sample of  $\sim 1700$  LBGs at  $z \sim 3 - 6$  from the GOODS-MUSIC catalog (Santini et al. 2009) with deep photometry from the  $U$  band to  $8 \mu\text{m}$ . Using a modified version of the *HyperZ* photometric redshift code which takes into account nebular emission (Schaerer & de Barros 2009), we explore a range of star formation history (constant, exponentially decreasing and rising). We explore the wide parameter space in redshift, metallicity, age, and extinction described by the Calzetti law (Calzetti et al. 2000), varying e-folding timescales for star formation, and whether or not nebular emission is included. The main physical parameters derived from our models are the stellar mass, age, reddening, star formation rate SFR and specific SFR. Furthermore our models also provide information on the characteristic SF timescale.

Our method and the selected sample has been described in Sects. 2 and 3. The detailed model results concerning the physical parameters, correlations among them, and describing the redshift evolution of the galaxy properties have been presented in Sect. 4. Our main results can be summarized as follows:

- Independently of the adopted star formation history, we find that at all redshifts  $\sim 65\%$  of the galaxies are better fitted with nebular emission and  $\sim 35\%$  without (Fig. 2). For galaxies with  $z \in [3.8, 5]$ , these two groups are clearly identified by their  $3.6\mu\text{m}$ - $4.5\mu\text{m}$  color (Fig. 3), which is correlated to strong  $H\alpha$  emission (cf. Shim et al. 2011). Furthermore, this color can not be reproduced if we impose an age limitation  $> 50\text{Myr}$  (Fig. 4). This observed color distribution clearly indicates the existence of galaxies with strong emission lines and others with few or no discernible signs of emission lines. Our SED modeling reveals the presence of two LBG groups (dubbed “strong” and “weak” emitters respectively) at all redshifts studied here ( $z \sim 3-6$ ), from U-drops to i-drops.
- Models including the effects of nebular emission and accounting for variable (declining or rising) star formation histories naturally separate the two LBG groups according to current star formation rate, the group of “strong” emitters showing a larger SFR than the “weak” emitters (Fig. 17). In a scenario of declining star formation histories these groups could be seen as starbursts and “post-starbursts”, with age differences compatible with this suggestion. Models with constant SFR and nebular emission cannot reproduce the observed range of  $3.6\mu\text{m}$ - $4.5\mu\text{m}$  colors of  $z \sim 3.8-5$  galaxies.
- Independently of the star formation history, the inclusion of nebular emission leads on average to significantly younger ages (Fig. 5), since nebular lines in optical (rest-frame) are able to mimic a Balmer break. This confirms our earlier findings (Schaerer & de Barros 2009, 2010) for larger samples and over a wider redshift range.
- We find that the derived dust attenuation mainly depends on the assumed star formation history, and that the treatment of nebular emission does not lead to a general systematic shift. The largest attenuation is found with rising star formation histories, since these always predict very recent star formation and hence UV emission, as already discussed e.g. by (Schaerer & Pelló 2005). In this case the inclusion of nebular emission decreases somewhat the average attenuation, whereas the attenuation increases for declining SFHs, and remains unchanged for constant SFR.
- Based on our SED fits for  $700 z \sim 4$  LBGs we propose a new average relation between the observed UV slope  $\beta$  and the attenuation  $A_V$ . Our relation deviates somewhat from the classical relation which assumes constant SFR and ages  $\geq 100$  Myr, and leads to a higher attenuation for a given  $\beta$  slope.
- Taking into account nebular emission, the stellar masses derived from the SED fits decrease by  $\sim 0.4$  on average, and by larger amounts ( $\sim 0.4 - 0.9$  dex) for LBGs from the “strong” emitter group (Fig. 11). We find a trend of increasing dust attenuation with stellar mass (Fig. 14) as already suggested earlier for  $z \sim 6-8$  galaxies (Schaerer & de Barros 2010), and a trend of increasing age with galaxy mass (Fig. 13).
- Given the large scatter found in the  $\text{SFR}-M_\odot$  relation for all models with variable star formation histories, we also find a large scatter for the specific star formation rate sSFR with stellar mass and at all redshifts. Our favored models show a higher median sSFR at  $z \sim 3 - 6$  than derived by previous studies (e.g. González et al. 2011). Our results tend to indicate an increase of the median sSFR with redshift, as advocated by several theoretical galaxy formation and evolution models (Bouché et al. 2010; Dutton et al. 2010; Weinmann et al. 2011)
- While uncertainties on  $\tau$  remain large, our SED fits favor short median star formation timescales ( $\lesssim 300$  Myr). Furthermore we find tentative evidence for decreasing values of  $\tau$  with decreasing UV luminosity among the sample of “strong” emitters.
- As already shown in Stark et al. (2009), constant star formation seems to be irreconcilable with the non-evolution of the  $M_\star - M_{1500}$  relation between  $z \sim 5$  to 3 and with the UV luminosity function.
- The rising average star formation of Finlator et al. (2011) used in this study cannot be representative of the typical history of many individual LBGs over long time since the predicted increase – if continuing into the future – is too fast. Indeed, this would lead to a strong increase of median stellar masses and the median SFR from one redshift to another, which would also imply an increase of dust attenuation to hide these objects from the LBG selection, since such very massive and strongly star-forming galaxies are not seen in the expected numbers.
- The presence of two groups of LBGs identified as active and more quiescent galaxies, our finding of general best SED fits with declining star formation histories, and the consistency of these results with constraints on duty cycles from clustering studies and other theoretical arguments (Lee et al. 2009a; Wyithe & Loeb 2011) all concur to consider episodic star formation as the scenario which best fits observations.

Our systematic and homogeneous analysis casts new light on the physical properties of LBGs from  $z \sim 3$  to 6 and possibly also to higher redshift (cf. Schaefer & de Barros 2010). Obviously, our results have a potentially important impact on a variety of questions, and several implications need to be worked out. On the other hand, our study also calls for new observations and tests.

For example, both our preferred models (variable star formation histories with nebular emission) imply a higher UV attenuation than what is currently derived, e.g. using the observed UV slope. At  $z \sim 4$  our models typically predict an increase by a factor  $\lesssim 3$ , but smaller changes at higher redshift. Implications on the cosmic star formation history and related topics will be worked out in a separate publication. If correct, a higher UV attenuation should lead to measurable changes e.g. in the IR emission of LBGs. A stacking analysis of the LBGs studied in this paper shows that the models presented here are all compatible with the current IR, submm, and radio observations (Greve et al.



2012, in preparation). Detailed predictions of the IR-mm emission from our galaxies are presented in Schaerer et al. (2012). Future, more sensitive observations with ALMA should be able to detect individual LBGs over a wide redshift range, to determine their attenuation, and hence also to distinguish different star formation histories (cf. Shim et al. 2011, Schaerer et al. 2012).

An important implication from our study is that the idea of simple, well-defined “star formation main sequence” with the majority of star-forming galaxies showing tight relation between SFR and  $M_*$ , suggested from other studies at lower redshift ( $z \sim 0-2$ , cf. Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007) may not be appropriate at high redshift ( $z \gtrsim 3$ ). Indeed a relatively small scatter is only obtained assuming constant star formation over long enough timescales ( $\gtrsim 50$  Myr), whereas our models clearly provide indications for variable star formation histories and episodic star formation (cf. Sects. 4.3.5, 4.2). A significant scatter is also obtained for models assuming rising or delayed star formation histories (see Sect. 4.3.7 and Schaerer et al. 2012), which are often suggested in the recent literature. If the scatter found from our models in the SFR– $M_*$  relation decreases with decreasing redshift, converging towards the results from other studies, remains to be seen. However, a smaller scatter is naturally found since constant star formation is usually *assumed* in most models or for establishing the standard SFR calibrations used in the literature (Daddi et al. 2007; Elbaz et al. 2007; González et al. 2011; Wuyts et al. 2011) or when analyzing stacked data, which naturally smoothes out any possible variation (Lee et al. 2010). Establishing more precisely the attenuation, current SFR, and star formation histories of galaxies is therefore crucial to shed more light on these questions.

Related to the above mentioned scatter is also the behavior of the specific star formation rate (sSFR) both with galaxy mass and on average with redshift. Basically the same questions and uncertainties concerning the SFR– $M_*$  relation apply here. In any case it must be recognized that the sSFR is strongly dependent on model assumptions, and seems to show a strong dependence on the galaxy mass. Whereas earlier determinations of the sSFR at  $z > 3$  were considered in conflict with recent galaxy evolution models (Bouché et al. 2010), our results are clearly in better agreement with the high sSFR values and the redshift evolution predicted by these models. A detailed confrontation of our results with such models and more refined ones, will hopefully provide further insight into galaxy formation and evolution models at high redshift.

Finally, it is clear that our study reveals new aspects on the possibly complex and variable star formation histories of high redshift galaxies. Whereas different arguments are found in the literature favouring e.g. short duty cycles and episodic star formation (cf. Sawicki & Yee 1998; Verma et al. 2007; Stark et al. 2009; Lee et al. 2009a; Wyithe & Loeb 2011) based on SED studies, on LBG clustering and other arguments, other studies favor long star formation timescales (Shapley et al. 2001; Lee et al. 2011; Shim et al. 2011). Our study uses for the first time quantitatively features probing emission lines, whose strength is naturally sensitive to relatively rapid variations of the recent SFR. That such variations have been more difficult to uncover before, may therefore not be surprising. Direct observations of the emission lines in high- $z$  LBGs should be very useful to test our models and provide more stringent constraints on the importance of nebular emission and on the star formation histories of distant galaxies (Schaerer et al. 2012). Other studies providing e.g. new measurement of LBG clustering at  $z > 4$ , searches for passive galaxies at high redshift, or theoretical studies on star

formation and regulation processes can also help improving our understanding of these important issues closely related to key questions on galaxy formation and evolution.

**Acknowledgements.** We would like to thank numerous colleagues who have contributed to interesting discussions and raised useful questions during the time where this work has been carried out. We thank them here collectively. We acknowledge the GOODS-MUSIC collaboration. The work of SdB and DS is supported by the Swiss National Science Foundation. DPS is supported by an STFC postdoctoral research fellowship. SdB wants to dedicate this work to the memory of Jean Faure-Brac.

## References

- Acquaviva, V., Vargas, C., Gawiser, E., & Guaita, L. 2011, ArXiv e-prints  
 Anders, P. & Fritze-v. Alvensleben, U. 2003, A&A, 401, 1063  
 Ando, M., Ohta, K., Iwata, I., et al. 2006, ApJ, 645, L9  
 Atek, H., Siana, B., Scarlata, C., et al. 2011, ApJ, 743, 121  
 Baker, A. J., Tacconi, L. J., Genzel, R., Lehnert, M. D., & Lutz, D. 2004, ApJ, 604, 125  
 Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, AJ, 132, 1729  
 Blanc, G. A., Adams, J. J., Gebhardt, K., et al. 2011, ApJ, 736, 31  
 Bolzonella, M., Miralles, J., & Pelló, R. 2000, A&A, 363, 476  
 Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001  
 Bouchet, P., Lequeux, J., Maurice, E., Prevot, L., & Prevot-Burnichon, M. L. 1985, A&A, 149, 330  
 Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Broadhurst, T. J., & Franx, M. 2004, ApJ, 611, L1  
 Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, ApJ, 705, 936  
 Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928  
 Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, ArXiv e-prints  
 Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151  
 Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000  
 Buat, V., Boissier, S., Burgarella, D., et al. 2008, A&A, 483, 107  
 Buat, V., Iglesias-Páramo, J., Seibert, M., et al. 2005, ApJ, 619, L51  
 Bunker, A. J., Stanway, E. R., Ellis, R. S., & McMahon, R. G. 2004, MNRAS, 355, 374  
 Burgarella, D., Le Floc’h, E., Takeuchi, T. T., et al. 2007, MNRAS, 380, 986  
 Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682  
 Castellano, M., Fontana, A., Grazian, A., et al. 2012, A&A, 540, A39  
 Charlot, S. & Longhetti, M. 2001, MNRAS, 323, 887  
 Cristiani, S., Appenzeller, I., Arnouts, S., et al. 2000, A&A, 359, 489  
 Curtis-Lake, E., McLure, R. J., Pearce, H. J., et al. 2012, MNRAS, 422, 1425  
 Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156  
 Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680  
 Doherty, M., Bunker, A. J., Ellis, R. S., & McCarthy, P. J. 2005, MNRAS, 361, 525  
 Domínguez, A., Siana, B., Henry, A. L., et al. 2012, ArXiv e-prints  
 Douglas, L. S., Bremer, M. N., Lehnert, M. D., Stanway, E. R., & Milvang-Jensen, B. 2010, MNRAS, 409, 1155  
 Dunne, L., Ivison, R. J., Maddox, S., et al. 2009, MNRAS, 394, 3  
 Dutton, A. A., van den Bosch, F. C., & Dekel, A. 2010, MNRAS, 405, 1690  
 Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33  
 Erb, D. K., Pettini, M., Shapley, A. E., et al. 2010, ApJ, 719, 1168  
 Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, ApJ, 644, 813  
 Eyles, L. P., Bunker, A. J., Ellis, R. S., et al. 2007, MNRAS, 374, 910  
 Eyles, L. P., Bunker, A. J., Stanway, E. R., et al. 2005, MNRAS, 364, 443  
 Ferguson, H. C., Dickinson, M., Giavalisco, M., et al. 2004, ApJ, 600, L107  
 Finkelstein, S. L., Rhoads, J. E., Malhotra, S., Grogin, N., & Wang, J. 2008, ApJ, 678, 655  
 Finlator, K., Davé, R., & Oppenheimer, B. D. 2007, MNRAS, 376, 1861  
 Finlator, K., Oppenheimer, B. D., & Davé, R. 2011, MNRAS, 410, 1703  
 Fioc, M. & Rocca-Volmerange, B. 1997, A&A, 326, 950  
 Giacomini, R., Zirm, A., Wang, J., et al. 2002, ApJS, 139, 369  
 Giavalisco, M., Dickinson, M., Ferguson, H. C., et al. 2004, ApJ, 600, L103  
 González, V., Labbé, I., Bouwens, R. J., et al. 2011, ApJ, 735, L34+  
 González, V., Labbé, I., Bouwens, R. J., et al. 2010, ApJ, 713, 115  
 Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, ApJ, 667, 79  
 Hathi, N. P., Malhotra, S., & Rhoads, J. E. 2008, ApJ, 673, 686  
 Hayes, M., Schaerer, D., Östlin, G., et al. 2011, ApJ, 730, 8  
 Hildebrandt, H., Arnouts, S., Capak, P., et al. 2010, A&A, 523, A31  
 Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 54  
 Kennicutt, Jr., R. C. 1998, ARA&A, 36, 189  
 Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. M. Iye & A. F. M. Moorwood, 1670–1681

- Le Fèvre, O., Vettolani, G., Paltani, S., et al. 2004, *A&A*, 428, 1043
- Lee, K.-S., Dey, A., Reddy, N., et al. 2011, *ApJ*, 733, 99
- Lee, K.-S., Giavalisco, M., Conroy, C., et al. 2009a, *ApJ*, 695, 368
- Lee, S.-K., Ferguson, H. C., Somerville, R. S., Wiklind, T., & Giavalisco, M. 2010, *ApJ*, 725, 1644
- Lee, S.-K., Idzi, R., Ferguson, H. C., et al. 2009b, *ApJS*, 184, 100
- Lidman, C., Hayes, M., Jones, D. H., et al. 2012, *MNRAS*, 420, 1946
- Madau, P. 1995, *ApJ*, 441, 18
- Madau, P., Pozzetti, L., & Dickinson, M. 1998, *ApJ*, 498, 106
- Maiolino, R., Nagao, T., Grazian, A., et al. 2008, *A&A*, 488, 463
- McLinden, E. M., Finkelstein, S. L., Rhoads, J. E., et al. 2011, *ApJ*, 730, 136
- McLure, R. J., Dunlop, J. S., de Ravel, L., et al. 2011, *MNRAS*, 418, 2074
- Mignoli, M., Cimatti, A., Zamorani, G., et al. 2005, *A&A*, 437, 883
- Nilsson, K. K., Möller, P., Möller, O., et al. 2007, *A&A*, 471, 71
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJ*, 660, L43
- Nonino, M., Dickinson, M., Rosati, P., et al. 2009, *ApJS*, 183, 244
- Oke, J. B. & Gunn, J. E. 1983, *ApJ*, 266, 713
- Ono, Y., Ouchi, M., Shimasaku, K., et al. 2010, *ApJ*, 724, 1524
- Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, *ApJS*, 176, 301
- Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, *ApJ*, 559, 620
- Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, *MNRAS*, 412, 1123
- Pentericci, L., Grazian, A., Fontana, A., et al. 2007, *A&A*, 471, 433
- Pettini, M., Kellogg, M., Steidel, C. C., et al. 1998, *ApJ*, 508, 539
- Popesso, P., Dickinson, M., Nonino, M., et al. 2009, *A&A*, 494, 443
- Prevot, M. L., Lequeux, J., Prevot, L., Maurice, E., & Rocca-Volmerange, B. 1984, *A&A*, 132, 389
- Ravikumar, C. D., Puech, M., Flores, H., et al. 2007, *A&A*, 465, 1099
- Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, *ArXiv e-prints*
- Reddy, N. A., Steidel, C. C., Fadda, D., et al. 2006, *ApJ*, 644, 792
- Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, *ApJS*, 175, 48
- Roche, N. D., Dunlop, J., Caputi, K. I., et al. 2006, *MNRAS*, 370, 74
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Santini, P., Fontana, A., Grazian, A., et al. 2009, *A&A*, 504, 751
- Sawicki, M. 2012, *MNRAS*, 421, 2187
- Sawicki, M. & Yee, H. K. C. 1998, *AJ*, 115, 1329
- Schaerer, D. & de Barros, S. 2009, *A&A*, 502, 423
- Schaerer, D. & de Barros, S. 2010, *A&A*, 515, A73+
- Schaerer, D. & de Barros, S. 2011, *ArXiv e-prints*
- Schaerer, D., de Barros, S., & Sklias, P. 2012, *ArXiv e-prints*
- Schaerer, D., de Barros, S., & Stark, D. P. 2011, *A&A*, 536, A72
- Schaerer, D. & Pelló, R. 2005, *MNRAS*, 362, 1054
- Schenker, M. A., Stark, D. P., Ellis, R. S., et al. 2012, *ApJ*, 744, 179
- Shapley, A. E., Steidel, C. C., Adelberger, K. L., et al. 2001, *ApJ*, 562, 95
- Shim, H., Chary, R.-R., Dickinson, M., et al. 2011, *ApJ*, 738, 69
- Siana, B., Teplitz, H. I., Chary, R.-R., Colbert, J., & Frayer, D. T. 2008, *ApJ*, 689, 59
- Stanway, E. R., Bunker, A. J., & McMahon, R. G. 2003, *MNRAS*, 342, 439
- Stark, D. P., Bunker, A. J., Ellis, R. S., Eyles, L. P., & Lacy, M. 2007, *ApJ*, 659, 84
- Stark, D. P., Ellis, R. S., Bunker, A., et al. 2009, *ApJ*, 697, 1493
- Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, *MNRAS*, 408, 1628
- Stringer, M., Cole, S., Frenk, C. S., & Stark, D. P. 2011, *MNRAS*, 414, 1927
- Strolger, L.-G., Riess, A. G., Dahlen, T., et al. 2004, *ApJ*, 613, 200
- Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, *ApJS*, 155, 271
- Teplitz, H. I., Desai, V., Armus, L., et al. 2007, *ApJ*, 659, 941
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898
- Trump, J. R., Impey, C. D., Kelly, B. C., et al. 2011, *ApJ*, 733, 60
- van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2005, *ApJ*, 631, 145
- van der Wel, A., Straughn, A. N., Rix, H.-W., et al. 2011, *ApJ*, 742, 111
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, 478, 83
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2005, *A&A*, 434, 53
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, *A&A*, 454, 423
- Verma, A., Lehnert, M. D., Förster Schreiber, N. M., Bremer, M. N., & Douglas, L. 2007, *MNRAS*, 377, 1024
- Weinmann, S. M., Neistein, E., & Dekel, A. 2011, *MNRAS*, 417, 2737
- Williams, R. E., Blacker, B., Dickinson, M., et al. 1996, *AJ*, 112, 1335
- Wolf, C., Meisenheimer, K., Kleinheinrich, M., et al. 2004, *A&A*, 421, 913
- Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011, *ApJ*, 738, 106
- Wuyts, S., Franx, M., Cox, T. J., et al. 2009, *ApJ*, 696, 348
- Wyithe, J. S. B. & Loeb, A. 2011, *MNRAS*, 413, L38
- Xu, C., Pirzkal, N., Malhotra, S., et al. 2007, *AJ*, 134, 169
- Yabe, K., Ohta, K., Iwata, I., et al. 2009, *ApJ*, 693, 507
- Yan, H., Dickinson, M., Giavalisco, M., et al. 2006, *ApJ*, 651, 24
- Yan, H., Dickinson, M., Stern, D., et al. 2005, *ApJ*, 634, 109
- Yang, Y., Flores, H., Hammer, F., et al. 2008, in *Astronomical Society of the Pacific Conference Series*, Vol. 390, Pathways Through an Eclectic Universe, ed. J. H. Knapen, T. J. Mahoney, & A. Vazdekis, 502–+
- Zackrisson, E., Bergvall, N., & Leitet, E. 2008, *ApJ*, 676, L9

## Appendix A:

**Table A.1.** Galaxies properties over  $3 \leq z \leq 6$  for constant star formation and solar metallicity set of models, in  $M_{1500}$  bins. For each parameter, we give the median value and 68% confidence limits derived from the probability distribution function. *Tables will be available in the published version.*

**Table A.2.** Same as Table A.1 for decreasing star formation history with variable timescale and metallicity  $Z$ .

**Table A.3.** Same as Table A.1 for rising star formation history with variable metallicity  $Z$ .